

FIGURE 1

CGGACGCGTGGGTGCGAGGCGAAGGTGACCGGGGACCGAGCATTTCAGATCTGCTCGGTAGA
CCTGGTGCACCACCACC**ATG**TTGGCTGCAAGGCTGGTGTGTCTCCGGACACTACCTTCTAGG
GTTTTCCACCCAGCTTTCACCAAGGCCTCCCCTGTTGTGAAGAATTCCATCACGAAGAATCA
ATGGCTGTTAACACCTAGCAGGGAATATGCCACCAAACAAGAATTGGGATCCGGCGTGGGA
GAACTGGCCAAGAAGCTCAAAGAGGCAGCATTGGAACCATCGATGGAAAAAATATTTAAATTT
GATCAGATGGGAAGATGGTTTGTGCTGGAGGGGCTGCTGTTGGTCTTGGAGCATTGTGCTA
CTATGGCTTGGGACTGTCTAATGAGATTGGAGCTATTGAAAAGGCTGTAATTTGGCCTCAGT
ATGTCAAGGATAGAATTCATTCCACCTATATGTACTTAGCAGGGAGTATTGGTTTAAACAGCT
TTGTCTGCCATAGCAATCAGCAGAACGCCTGTTCTCATGAACCTTCATGATGAGAGGCTCTTG
GGTGACAATTGGTGTGACCTTTGCAGCCATGGTGGAGCTGGAATGCTGGTACGATCAATAC
CATATGACCAGAGCCCAGGCCCAAAGCATCTTGCTTGGTTGCTACATTCTGGTGTGATGGGT
GCAGTGGTGGCTCCTCTGACAATATTAGGGGGTCTCTTCTCATCAGAGCTGCATGGTACAC
AGCTGGCATTGTGGGAGGCCTCTCCACTGTGGCCATGTGTGCGCCCAGTGAAAAGTTTCTGA
ACATGGGTGCACCCCTGGGAGTGGGCCCTGGGTCTCGTCTTTGTGTCTCATTGGGATCTATG
TTTCTTCCACCTACCACCGTGGCTGGTGCCACTCTTTACTCAGTGGCAATGTACGGTGGATT
AGTTCTTTTTCAGCATGTTCTTCTGTATGATACCCAGAAAGTAATCAAGCGTGCAGAAGTAT
CACCAATGTATGGAGTTCAAAAATATGATCCCATTAACCTCGATGCTGAGTATCTACATGGAT
ACATTAAATATATTTATGCGAGTTGCAACTATGCTGGCAACTGGAGGCAACAGAAAGAAAT**AG**
AAGTGACTCAGCTTCTGGCTTCTCTGCTACATCAAATATCTTGTTTAAATGGGGCAGATATGC
ATTAAATAGTTTGTACAAGCAGCTTTCGTTGAAGTTTAGAAGATAAGAAACATGTCATCATA
TTTAAATGTTCCGGTAATGTGATGCCTCAGGTCTGCCTTTTTTTCTGGAGAATAAATGCAGT
AATCCTCTCCCAAATAAGCACACACATTTTCAATTCTCATGTTTGAGTGATTTTAAATGTT
TTGGTGAATGTGAAAACATAAGTTTGTGTGTCATGAGAATGTAAGTCTTTTTTCTACTTTAAAA
TTTAGTAGGTTCACTGAGTAACTAAAATTTAGCAAACCTGTGTGTCATATTTTTTTGGAGT
GCAGAAATATTGTAATTAATGTCATAAGTGATTTGGAGCTTTGGTAAAGGGACCAGAGAGAAG
GAGTCACCTGCAGTCTTTTGTTTTTTTAAATACTTAGAACTTAGCACTTGTGTTATTGATTA
GTGAGGAGCCAGTAAGAAACATCTGGGTATTTGGAAACAAGTGGTCATTGTTACATTCATTT
GCTGAACCTTAACAAAACCTGTTTCATCCTGAAACAGGCACAGGTGATGCATTCTCCTGCTGTTG
CTTCTCAGTGCTCTCTTTCCAATATAGATGTGGTCATGTTTGAAGTTGTACAGAATGTTAATC
ATACAGAGAATCCTTGATGGAATTATATATGTGTGTTTACTTTTGAATGTTACAAAAGGAA
ATAACTTTAAACACTATTCTCAAGAGAAAATATTCAAAGCATGAAATATGTTGCTTTTTCCAG
AATACAAACAGTATACTCATG

CGGACGCGTGGGTGCGAGGCGAAGGTGACCGGGGACCGAGCATTTCAGATCTGCTCGGTAGA

FIGURE 2

MLAARLVCLRTLPSRVFHPAFTKASPVVKNSITKNQWLLTPSREYATKTRIGIRRGRTGQEL
KEAALEPSMEKIFKIDQMGRWVAVAGGAAGVGLGALCYGGLSNEIGAIEKAVIWPQYVKDRI
HSTYMYLAGSIGLTALSAIAISRTPVLMNFMMRGSWVTIGVTFAAMVGAGMLVRSIPYDQSP
GPKHLAWLLHSGVMGAVVAPLTILGGPLLIRAAWYTAGIVGGLSTVAMCAPSEKFLNMGAPL
GVGLGLVFVSSLGSMFLPPTTVAGATLYSVAMYGGLVLFMSMFLLYDTQKVIKRAEVSPMYGV
QKYDPINSMLSIYMDTLNIFMRVATMLATGGNRKK

Sequence

FIGURE 3

GAAGGCTGCCTCGCTGGTCCGAATTCGGTGGCGCCACGTCCGCCCCGTCTCCGCGCTTCTGCAT
CGCGGCTTCGGCGGCTTCCACCTAGACACCTAACAGTCGCGGAGCCGGCCGCGTCGTGAGGG
GGTCGGCACGGGGAGTCGGGCGGTCTTGTGCATCTTGGCTACCTGTGGGTCTGAAG**ATGT**CGG
ACATCGGAGACTGGTTCAGGAGCATCCCGGCGATCACGCGCTATTGGTTCGCCGCCACCGTC
GCCGTGCCCTTGGTCGGCAAACCTCGGCCTCATCAGCCCGGCCTACCTCTTCCCTCTGGCCCGA
AGCCTTCCTTTATCGCTTTCAGATTTGGAGGCCAATCACTGCCACCTTTTATTTCCCTGTGG
GTCCAGGAACCTGGATTTCTTTATTTGGTCAATTTATATTTCTTATATCAGTATTCTACGCGA
GATACCCCAATGGACTTGGGAGGAAGAAATTTTCTATCCACACCTCAGTTTTTGTACCGCTGG
GATTTGCATCGTGATTACTGGCTTAGCAATGGATATGCAGTTGCTGATGATTCTCTGATCA
TGTCAGTACTTTATGTCTGGGCCCAGCTGAACAGAGACATGATTGTATCATTTTTGGTTTTGGA
ACACGATTTAAGGCCTGCTATTTACCCTGGGTATCCTTGGATTCAACTATATCATCGGAGG
CTCGGTAATCAATGAGCTTATTGGAAATCTGGTTGGACATCTTTATTTTTTCTAATGTTCA
GATACCCCAATGGACTTGGGAGGAAGAAATTTTCTATCCACACCTCAGTTTTTGTACCGCTGG
CTGCCCAGTAGGAGAGGAGGAGTATCAGGATTTGGTGTGCCCCCTGCTAGCATGAGGCGAGC
TGCTGATCAGAATGGCGGAGGCGGGAGACACAACCTGGGGCCAGGGCTTTCGACTTGGAGACC
AG**TGA**AGGGGGCGGCCTCGGGCAGCCGCTCCTCTCAAGCCACATTTCCCTCCCAGTGCTGGGTG
CACTTAACAACCTGCGTTCTGGCTAACACTGTTGGACCTGACCCACACTGAATGTAGTCTTTT
AGTACGAGACAAAAGTTTCTTAAATCCCGAAGAAAAATATAAGTGTTCACAAAGTTTCACGAT
TCTCATTTCAAGTCCTTACTGCTGTGAAGAACAAATACCAACTGTGCAAATTGCAAAACTGAC
TACATTTTTTTGGTGTCTTCTCTTCTCCCCTTTCCGTCTGAATAATGGGTTTTAGCGGGTCCCT
AATCTGCTGGCATTGAGCTGGGGCTGGGTACCAAAACCTTCCCAAAGGACCTTATCTCTT
TCTTGACACATGCCTCTCTCCCACCTTTCCCAACCCCCACATTTGCAACTAGAAAAAGTTG
CCATAAAATTTGCTCTGCCCCTTGACAGGTTCTGTTATTTATTGACTTTTGCCAAGGCTGGTC
ACAACAATCATATTCACGTTATTTTCCCCTTTTGGTGGCAGAACTGTTACCAATAGGGGGAG
AAGACAGCCACGGATGAAGCGTTTCTCAGCTTTTGGAAATTGCTTCGACTGACATCCGTTGTT
AACCGTTTGCCACTCTTCAGATATTTTTTATAAAAAAAGTACCCTGAGTTCATGAGGGCCA
CAGATTGGTTATTAATGAGATACGAGGGTTGGTGTCTGGGTGTTTGTCTTCTGAGCTAAGTGA
TCAAGACTGTAGTGGAGTTGCAGCTAACATGGGTAGGTTTAAACCATGGGGGATGCACCCC
TTTGCGTTTCATATGTAGCCCTACTGGCTTTGTGTAGCTGGAGTAGTTGGGTTGCTTTGTGT
TAGGAGGATCCAGATCATGTTGGCTACAGGGAGATGCTCTCTTTGAGAGGTCCTGGGCATTG
ATTCCCATTTCATCTCATTTCTGGATATGTGTTTCATTGAGTAAAGGAGGAGAGACCCTCATA
CGCTATTTAAATGTCACTTTTTTGCCTATCCCCCGTTTTTTGGTCATGTTTCAATTAATTGT
GAGGAAGGCGCAGCTCCTCTCTGCACGTAGATCATTTTTTAAAGCTAATGTAAGCACATCTA
AGGGAATAACATGATTTAAGGTTGAAATGGCTTTAGAATCATTTGGGTTTGAGGGTGTGTTA
TTTTGAGTCATGAATGTACAAGCTCTGTGAATCAGACCAGCTTAAATACCCACACCTTTTTT
TCGTAGGTGGGCTTTTCTATCAGAGCTTGGCTCATAACCAAATAAAGTTTTTTGAAGGCCA
TGGCTTTTACACAGTTATTTTATTTATGACGTTATCTGAAAGCAGACTGTTAGGAGCAGT
ATTGAGTGGCTGTACACTTTGAGGCAACTAAAAAGGCTTCAAACGTTTTGATCAGTTTCTT
TTCAGGAAACATTGTGCTCTAACAGTATGACTATTCTTTCCCCCACTCTTAAACAGTGTGAT
GTGTGTTATCCTAGGAAATGAGAGTTGGCAAACAACCTTCTCATTTTGAATAGAGTTTGTGTG
TACTTCTCCATATTTAATTTATATGATAAAATAGGTGGGGAGAGTCTGAACCTTAACTGTCA
TGTTTTGTGTTTCATCTGTGGCCACAATAAAGTTTACTTGTAAAATTTTAGAGGCCATTACT
CCAATTATGTTGCACGTACACTCATTTGTACAGGCGTGGAGACTCATTGTATGTATAAGAATA
TTTCTGACAGTGAGTGACCCGGAGTCTCTGGTGTACCTCTTACCAGTCAGCTGCCTGCGAG
CAGTCATTTTTTCTAAAGGTTTACAAGTATTTAGAACTTTTCAGTTTCAGGGCAAAATGTTT
ATGAAGTTATTCCCTCTTAAACATGGTTAGGAAGCTGATGACGTTATTGATTTTTGTCTGGATT
ATGTTTCTGGAATAATTTTACCAAAACAAGCTATTTGAGTTTTGACTTGACAAGGCAAAACA
TGACAGTGGATTCTCTTTACAAATGGAATAAAAAAATCCTTATTTTGTATAAAGGACTTCCC
TTTTTGTAACATAATCCTTTTTTATTGGTAAAAATTGTAAATTAAATGTGCAACTTG

MSDIGDWFRSIPAITRYWFAATVAVPLVGKLG LISPAYLFLWPEAFLYRFQIWRPITATFYF
PVGPGTGFLYLVNLYFLYQYSTRLETGAFGDRPADYLFMLLFNWICIVITGLAMDMQLLMIP
LIMSVLYVWAQLNRDMIVSFWFGTRFKACYLPWVILGFNYIIGGSVINELIGNLVGHLYFFL
MFRYPMDLGGRNFLSTPQFLYRWLPSRRGGVSGFGVPPASMRRAADQNGGGGRHNWGQGFRLL
GDQ

amino acids 98-116, 152-172

amino acids 89-95, 168-174, 176-182, 215-221, 221-227, 237-243

amino acids 218-222

FIGURE 5

GGGGCCGCGGTCTAGGGCGGCTACGTGTGTTGCCATAGCGACCATTTTGCATTAACTGGTTG
GTAGCTTCTATCCTGGGGGCTGAGCGACTGCGGGCCAGCTCTTCCCCTACTCCCTCTCGGCT
CCTTGTGGCCCAAAGGCCTAACCGGGGTCCGGCGGTCTGGCCTAGGGATCTTCCCCGTTGCC
CCTTTGGGGCGGG**ATG**GCTGCGGAAGAAGAAGACGAGGTGGAGTGGGTAGTGGAGAGCATCG
CGGGGTTCTGCGAGGCCCAGACTGGTCCATCCCCATCTTGGACTTTGTGGAACAGAAATGT
GAAGTTAACTGCAAAGGAGGGCATGTGATAACTCCAGGAAGCCCAGAGCCGGTGATTTTGGT
GGCCTGTGTTCCCCTTGTTTTTGATGATGAAGAAGAAAGCAAATTGACCTATACAGAGATTC
ATCAGGAATACAAAGAACTAGTTGAAAAGCTGTTAGAAGGTTACCTCAAAGAAATTGGAATT
AATGAAGATCAATTTCAAGAAGCATGCACTTCTCCTCTTGCAAAGACCCATACATCACAGGC
CATTTTGCAACCTGTGTTGGCAGCAGAAGATTTTACTATCTTTAAAGCAATGATGGTCCAGA
AAAACATTGAAATGCAGCTGCAAGCCATTCGAATAATTCAAGAGAGAAATGGTGTATTACCT
GACTGCTTAACCGATGGCTCTGATGTGGTCAGTGACCTTGAACACGAAGAGATGAAAATCCT
GAGGGAAGTTCTTAGAAAAATCAAAGAGGAATATGACCAGGAAGAAGAAAGGAAGAGGAAAA
AACAGTTATCAGAGGCTAAACAGAAGAGCCACAGTGCAATTCAGTGAAGCTGCAATAATG
AATAATTCCCAAGGGGATGGTGAACATTTTGCACACCCACCCTCAGAAGTTAAATGCATTT
TGCTAATCAGTCAATAGAACCTTTGGGAAGAAAAGTGGAAAGGTCTGAAACTTCCTCCCTCC
CACAAAAGGCCTGAAGATTCCTGGCTTAGAGCATGCGAGCATTGAAGGACCAATAGCAAAC
TTATCAGTACTTGGAACAGAAGAACTTCGGCAACGAGAACACTATCTCAAGCAGAAGAGAGA
TAAGTTGATGTCCATGAGAAAGGATATGAGGACTAAACAGATACAAAATATGGAGCAGAAAG
GAAAACCCACTGGGGAGGTAGAGGAAATGACAGAGAAACCAGAAATGACAGCAGAGGAGAAG
CAAACATTACTAAAGAGGAGATTGCTTGCAGAGAACTCAAAGAAGAAGTTATTAATAAG**TA**
ATAATTAAGAACAATTTAACAAAATGGAAGTTCAAATTGTCTTAAAAATAAATTATTTAGTC
CTTACACTG

FIGURE 6

MAAEEDEVEWVVESIAGFLRGPDWSIPILDFVEQKCEVNCKGGHVITPGSPEPVILVACVP
LVFDDEEESKLTYTEIHQEKELVEKLLEGYLKEIGINEDQFQEAactsPLAKTHTSQAILQP
VLAAEDFTIFKAMMVQKNIEMQLQAIRIIQERNGLVPDCLTDGSDVVSdleHEEMKILREVL
RKSKEEYDQEEERKRKKQLSEAKTEEPTVHSSEAAIMNNSQGDGEHFAHPPSEVKMHFANQS
IEPLGRKVERSETSSLPQKGLKIPGLEHASIEGPIANLSVLGTEELRQREHYLKQKRDKLMS
MRKDMRTKQIQNMEQKGKPTGEVEEMTEKPEMTAEKQTLLKRLLAEKLKEEVINK

N-glycosylation sites.

amino acids 224-228, 246-250, 285-289

N-myristoylation site.

amino acids 273-279

Amidation site.

amino acids 252-256

Cytosolic fatty-acid binding proteins.

amino acids 78-108

FIGURE 7

GGGCACAGCACATGTGAAGTTTTTGATGATGAAGAAGAAAGCAAATTGACCTATACAGAGAT
TCATCAGGAATACAAAGAACTAGTTGAAAAGCTGTTAGAAGGTTACCTCAAAGAAATTGGAA
TTAATGAAGATCAATTTCAAGAAGCATGCACTTCTCCTCTTGCAAAGACCCATACATCACAG
GCCATTTTTGCAACCTGTGTTGGCAGCAGAAGATTTTACTATCTTTAAAGCAATGATGGTCC
AGAAAAACATTGAAATGCAGCTGCAAGCCATTCGAATAATTCAAGAGAGAAATGGTGTATTA
CCTGACTGCTTAACCGATGGCTCTGATGTGGTCAGTGACCTTGAACACGAAGAGATGAAAAT
CCTGAGGGAAGTTCTTAGAAAATCAAAAGAGGAATATGACCAGGAA

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189

FIGURE 8

GCGTGGTTTTTGTCTGCAATAGGCGGCTTAGAGGGAGGGGCTTTTTTCGCCTATACCTACTG
 TAGCTTCTCCACGTATGGACCCTAAAGGCTACTGCTGCTACTACGGGGCTAGACAGTTACTG
 TCTCAGCTCTAGGATGTGCGTTCTTCCACTAGAAGCTCTTCTGAGGGAGGTAATTA AAAAAC
 AGTGGAA**ATG**GAAAAACAGTGCTGTAGTCATCCTGTAATATGCTCCTTGTCAACAATGTATAC
 ATTCCTGCTAGGTGCCATATTCATTGCTTTAAGCTCAAGTCGCATCTTACTAGTGAAGTATT
 CTGCCAATGAAGAAAAACAAGTATGATTATCTTCCAACACTGTGAATGTGTGCTCAGAACTG
 GTGAAGCTAGTTTTCTGTGTGCTTGTGTCATTCTGTGTTATAAAGAAAGATCATCAAAGTAG
 AAATTTGAAATATGCTTCCTGGAAGGAATTCTCTGATTTTCATGAAGTGGTCCATTCCCTGCCT
 TTCTTTATTTTCCTGGATAACTTGATTGTCTTCTATGTCTGTCTATCTTCAACCAGCCATG
 GCTGTTATCTTCTCAAAATTTTAGCATTATAACAACAGCTCTTCTATTCAAGGATAGTGTGAA
 GAGGCGTCTAAACTGGATCCAGTGGGCTTCCCTCCTGACTTTATTTTTGTCTATTGTGGCCT
 TGACTGCCGGGACTAAAACCTTTACAGCACAACCTTGGCAGGACGTGGATTTTCATCACGATGCC
 TTTTTTCAGCCCTTCCAATTCCTGCCTTCTTTTCAGAAGTGAGTGTCCAGAAAAGACAATTG
 TACAGCAAAGGAATGGACTTTTCCTGAAGCTAAATGGAACACCACAGCCAGAGTTTTTCAGTC
 ACATCCGTCTTGGCATGGGCCATGTTCTTATTATAGTCCAGTGTTTTATTTCTTCAATGGCT
 AATATCTATAATGAAAAGATACTGAAGGAGGGGAACCAGCTCACTGAAAGCATCTTCATACA
 GAACAGCAAACCTCTATTTCTTTGGCATTCGTTTAATGGGCTGACTCTGGGCCTTCAGAGGA
 GTAACCGTGATCAGATTAAGAAGTGTGGATTTTTTTATGGCCACAGTGCATTTTCAGTAGCC
 CTTATTTTTGTAACTGCATTCCAGGGCCTTTCAGTGGCTTTCATTCTGAAGTTCCTGGATAA
 CATGTTCCATGTCTTGATGGCCCAGGTTACCACTGTCATTATCACAAAGTGTCTGTCTCTGG
 TCTTTGACTTCAGGCCCTCCCTGGAATTTTTCTTGGAAGCCCCATCAGTCCTTCTCTCTATA
 TTTATTTTATAATGCCAGCAAGCCTCAAGTTCGGAATACGCACCTAGGCAAGAAAGGATCCG
 AGATCTAAGTGGCAATCTTTGGGAGCGTTCCAGTGGGGATGGAGAAGAACTAGAAAGACTTA
 CCAAACCCAAGAGTGATGAGTCAGATGAAGATACTTTCT**TAA**CTGGTACCCACATAGTTTGCA
 GCTCTCTTGAACCTTATTTTCACATTTTCAGTGTTTGTAATATTTATCTTTTCACTTTGATA
 AACCAGAAATGTTTCTAAATCCTAATATTCTTTGCATATATCTAGCTACTCCCTAAATGGTT
 CCATCCAAGGCTTAGAGTACCCAAAGGCTAAGAAATCTAAAGAACTGATACAGGAGTAACA
 ATATGAAGAATTCATTAATATCTCAGTACTTGATAAATCAGAAAGTTATATGTGCAGATTAT
 TTTCTTGGCCTTCAAGCTTCCAAAAAAGTGTAAATAATCATGTTAGCTATAGCTTGTATAT
 ACACATAGAGATCAATTTGCCAAATATTCACAATCATGTAGTTCTAGTTTACATGCCAAAGT
 CTTCCCTTTTTAACATTATAAAAGCTAGGTTGTCTCTTGAATTTTGGAGGCCCTAGAGATAGT
 CATTTTGCAAGTAAAGAGCAACGGGACCCCTTCTAAAAACGTTGGTTGAAGGACCTAAATAC
 CTGGCCATACCATAGATTTGGGATGATGTAGTCTGTGCTAAATATTTTGCTGAAGAAGCAGT
 TTCTCAGACACAACATCTCAGAATTTTAATTTTTAGAAATTCATGGGAAATTGGATTTTTGT
 AATAATCTTTTGATGTTTTAAACATTGGTTCCCTAGTCAACCATAGTTACCACTTGTATTTTA
 AGTCATTTAAACAAGCCACGGTGGGGCTTTTTTCTCCTCAGTTTGAGGAGAAAAATCTTGAT
 GTCATTACTCCTGAATTATTACATTTTGGAGAATAAGAGGGCATTTTATTTTATTAGTTACT
 AATTCAAGCTGTGACTATTGTATATCTTTCCAAGAGTTGAAATGCTGGCTTCAGAATCATAC
 CAGATTGTGCTAGTGAAGCTGATGCCTAGGAACCTTTTAAAGGGATCCTTTCAAAGGATCACTT
 AGCAAACACATGTTGACTTTTAACTGATGTATGAATATTAATACTCTAAAAATAGAAAGACC
 AGTAATATATAAGTCACTTTACAGTGCTACTTCACACTTAAAGTGCATGGTATTTTTTCATG
 GTATTTTGCATGCAGCCAGTTAACTCTCGTAGATAGAGAAGTCAGGTGATAGATGATATTAA
 AAATTAGCAAACAAAAGTGACTTGCTCAGGGTCATGCAGCTGGGTGATGATAGAAGAGTGGG
 CTTTAACTGGCAGGCCTGTATGTTTACAGACTACCATACTGTAAATATGAGCTTTATGGTGT
 CATTCTCAGAACTTATACATTTCTGCTCTCCTTTCTCCTAAGTTTCATGCAGATGAATATA
 AGGTAATATACTATTATATAATTCATTTGTGATATCCACAATAATATGACTGGCAAGAATTG
 GTGGAAATTTGTAATTAAATAATTATTAAACCT

FIGURE 9

MEKQCCSHPVICSLSTMYTFLLGAIFIALSSSRILLVKYSANEENKYDYLPTTVNVCSELVK
LVFCVLVSFCVIKKDHQSRNLKYASWKEFSDFMKWSIPAFLYFLDNLIVFYVLSYLQPAMAV
IFS NFSIITTALLFRIVLKRRNLNWIQWASLLTLFLSIVALTAGTKTLQHNLAGRGFHHDAFF
SPSNSCLLFRSECPRKDNCTAKEWTFPEAKWNTTARVFSHIRLGMGHVLIIVQCFISSMANI
YNEKILKEGNQLTESIFIQNSKLYFFGILENGLTLGLQRSNRDQIKNCGFFYGHSAFSVALI
FVTAFQGLSVAFILKFLDNMFHVLMAQVTTVIITT VSVLVFDFRPSLEFFLEAPSVLLSIFI
YNASKPQVPEYAPRQERIRDL SGNLWERSSSGDGEELERLTKPKSDESDETF

Transmembrane domains:

amino acids 16-36 (type II), 50-74, 147-168, 229-250, 271-293,
298-318, 328-368

N-glycosylation sites.

amino acids 128-132, 204-208, 218-222, 374-378

Glycosaminoglycan attachment site.

amino acids 402-406

N-myristoylation sites.

amino acids 257-263, 275-281, 280-286, 284-290, 317-323

FIGURE 10

CGTGCCTGCGCAATGGGTGTCGGGTCCGCTTTTTCCCAATCCGGACGTAATCGTGGTTTTTG
TTCTGCAATAGGCGGCTTAGAGGGAGGGGCTTTTTCGCCTATACCTACTGTAGCTTCTCCAC
GTATGGACCCTAAAGGCTACTGCTGCTACTACGGGGCTAGACAGTTACTGTCTCAGCTCTAG
GATGTGCGTTCTTCCACTAGAAGCTCTTCTGAGGGAGGTAATTAAAAACAGTGGAATGGAA
AAACAGTGCTGTAGTCATCCTGTAATATGCTCCTTGTCACAATGTATACATTCCTGCTAGG
TGCCATATTCATTGCTTTAAGCTCAAGTCGCATCTTACTAGTGAAGTATTCTGCCAATGAAG
AAAACAAGTATGATTATCTTCCAACACTGTGAATGTGTGCTCAGAACTGGTGAAGCTAGTT
TTCTGTGTGCTTGTGTCAATTCTGTGTTATAAAGAAAGATCATCAAAGTAGAAATTTGAAATA
TGCTTCCTGGAAGGAATTCTCTGATTCATGAAGTGGTCCATTCCTGCCTTTCTTTATTTCC
TGGATAACTTGATTGTCTTCTATGTCCTGTCCTATCTTCAACCAGCCATGGCTGTTATCTTC
TCAAATTTTAGCATTATAACAACAGCTCTTCTATTACAGGATAGTGCTGAAGAGGCGTCTAAA
CTGGATCCAGTGGGCTTCCCTCCTGACTTTATTTTTGTCTATTGTGGCCTTGACTGCCGGGA
CTAAAACCTTTA

CGGACGCGTGGGCGGACGCGTGGGGCCGGCTTGGCTAGCGCGCGGCGGCC
GTGGCTAAGGCTGTCTACGAAGCGAGCTTGGGAGGAGCAGCGGCCTGCGGGGCAGAGGAGCAT
CCCGTCTACCAAGTCCCAAGCGGCGTGGCCGCGGGTTCATGGCCAAAGGAGAAGGCGCCGAG
AGCGGCTCCGCGGCGGGGCTGCTACCCACCAGCATCCTCCAAAGCACTGAACGCCCGGCCCA
GGTGAAGAAAGAACCGAAAAAGAAACAACAGTTGTCTGTTTGAACAAGCTTTGCTATG
CACTTGGGGGAGCCCCCTACCAGGTGACGGGCTGTGCCCTGGGTTTCTTCCTTCAGATCTAC
CTATTGG**ATG**TGGCTCAGGTGGGCCCTTTCTCTGCCTCCATCATCCTGTTTGTGGGCCGAGC
CTGGGATGCCATCACAGACCCCTGGTGGGCCTCTGCATCAGCAAATCCCCCTGGACCTGCC
TGGGTGCGCTTATGCCCTGGATCATCTTCTCCACGCCCTGGCCGTCAATTGCCTACTTCCCT
ATCTGGTTCGTGCGGACTTCCCACACGGCCAGACCTATTGGTACCTGCTTTTCTATTGCCT
CTTTGAAACAATGGTCACGTGTTTCCATGTTCCCTACTCGGCTCTCACCATGTTTCATCAGCA
ACCGAGCAGACTGAGCGGGATTCTGCCACCGCCTATCGGATGACTGTGGAAGTGCTGGGCAC
AGTGCTGGGCACGGCGATCCAGGGACAAATCGTGGGCCAAGCAGACACGCCTTGTTTCCAGG
ACTTCAATAGCTCTACAGTAGCTTCACAAAGTGCCAACCATAACATGGCACCACTTCACAC
AGGGAACCGCAAAGGCATACCTGCTGGCAGCGGGGGTCAATTGTCTGTATCTATATAATCTG
TGCTGTCTCCTGATCCTGGGCGTGCGGGAGCAGAGAGAACCCTATGAAGCCCAGCAGTCTG
AGCCAATCGCCTACTTCCGGGGCCTACGGCTGGTTCATGAGCCACGGCCCATACATCAAACCT
ATTACTGGCTTCTCTTACCTCCTTGGCTTTTCATGCTGGTGGAGGGGAACCTTGTCTTGT
TTGCACCTACACCTTGGGCTTCCGCAATGAATTCCAGAATCTACTCCTGGCCATCATGCTCT
CGGCCACTTTAACCATTCCCATCTGGCAGTGGTTCTTGACCCGGTTTGGCAAGAAGACAGCT
GTATATGTTGGGATCTCATCAGCAGTGCCATTTCTCATCTTGGTGGCCCTCATGGAGAGTAA
CCTCATCATTAACATATGCGGTAGCTGTGGCAGCTGGCATCAGTGTGGCAGCTGCCTTCTTAC
TACCCTGGTCCATGCTGCCTGATGTCAATTGACGACTTCCATCTGAAGCAGCCCCACTTCCAT
GGAACCGAGCCCATCTTCTTCTCCTTCTATGTCTTCTTACCAAGTTTGCCTCTGGAGTGT
ACTGGGCATTTCTACCCTCAGTCTGGACTTTGCAGGGTACCAGACCCGTGGCTGCTCGCAGC
CGGAACGTGTCAAGTTTAACTGAACATGCTCGTGACCATGGCTCCCATAGTTCTCATCCTG
CTGGGCCTGCTGCTCTTCAAATGTACCCCATTGATGAGGAGAGGCGGCGGCAGAAATAAGAA
GGCCCTGCAGGCACTGAGGGACGAGGCCAGCAGCTCTGGCTGCTCAGAAACAGACTCCACAG
AGCTGGCTAGCATCCTCT**TAG**GGGCGGCCACGTTGCCCGAAGCCACCATGCAGAAGGCCACAG
AAGGGATCAGGACCTGTCTGCCGGCTTGTGAGCAGCTGGACTGCAGGTGCTAGGAAGGGAA
CTGAAGACTCAAGGAGGTGGCCCAGGACACTTGTGTGCTCACTGTGGGGCCGGCTGCTCTG
TGGCCTCCTGCCTCCCCTCTGCCTGCCTGTGGGGCCAAGCCCTGGGGCTGCCACTGTGAATA
TGCCAAGGACTGATCGGGCCTAGCCCGGAACACTAATGTAGAAACCTTTTTTTTACAGAGCC
TAATTAATAACTTAATGACTGTGTACATAGCAATGTGTGTGTATGTATATGTCTGTGAGCTA
TTAATGTTATTAATTTTCATAAAAGCTGGAAAGC

FIGURE 12

MWLRWALSLPPSSCLWAEFGMPQSQTPWWASASANPPGPAWVALCPGSSSPRPWPSLPTSSSG
SCPTSHTARPIGTCTFSIASLKQWSRVSMFPTRLSPCSSATEQTERDSATAYRMTVEVLGTVL
GTAIQGQIVGQADTPCFQDFNSSTVASQSANHTHGTTSHRETQKAYLLAAGVIVCIYIICAV
ILILGVREQREPYEAAQQSEPIAYFRGLRLVM SHGPYIKLITGFLFTSLAFMLVEGNFVLFCT
YTLGFRNEFQNLLLAIMLSATLTIPIWQWFLTRFGKKTAVYVGISSAVPFLILVALMESNLI
ITYAVAVAAGISVAAAFLLPWSMLPDVIDDFHLKQPHFHGTEPIFFSFYVFFTKFASGVSLG
ISTLSLDFAGYQTRGCSQPERVKFTLNMLVTMAPIVLILLGLLLFKMPIDEERRRQNKAL
QALRDEASSSGCSETDSTELASIL

Q
E
R
R
R
Q
N
K
A
L
I
D
E
A
S
S
S
G
C
S
E
T
D
S
T
E
L
A
S
I
L

FIGURE 13

GGGAAACGCAAAGGCATACCTGCTGGCAGCGGGGGTCATTGTCTGTATCTATATAATCTGT
GCTGTCATCCTGATCCTGGGCGTGCGGGAGCAGAGAGAACCCTATGAAGCCCAGCAGTCTGA
GCCAATCGCCTACTTCCGGGGCCTACGGCTGGTCATGAGCCACGGCCCATACATCAAACCTTA
TACTGGCTTCCTCTTCACCTCCTTGGCTTTCATGCTGGTGGAGGGGAACCTTGTCTTGTTT
TGCACCTACACCTTGGGCTTCCGCAATGAATTCCAGAATCTACTCCTGGCCATCATGCTCTC
GGCCACTTTAACCATTCCCATCTGGCAGTGGTTCTTGACCCGGTTTGGCAAGAAGACAGCTG
TATATGTTGGGATCTCATCAGCAGTGCCATTTCTCATCTTGGTGGCCCTCATGGAGAGTAAC
CTCATCATTACATATGCGGTAGCTGTGGCAGCTGGCATCAGTGTGGCAGCTGCCTTCTTACT
ACCTTGGTCCATGCTGCCTGATGTCATTGACGACTTCCATCTGAAGCAGCCCCACTTCCATG
GAACCGAGCCCAT

13/330

FIGURE 14

GGGGCTTCGGCGCCAGCGGCCAGCGCTAGTCGGTCTGGTAAGGATTTACAAAAGGTGCAGGT
ATGAGCAGGTCTGAAGACTAACATTTTGTGAAGTTGTAAAACAGAAAACCTGTTAGAAATGT
GGTGGTTTCAGCAAGGCCTCAGTTTCCTTCCTTCAGCCCTTGTAATTTGGACATCTGCTGCT
TTCATATTTTCATACATTACTGCAGTAACACTCCACCATATAGACCCGGCTTTACCTTATAT
CAGTGACACTGGTACAGTAGCTCCAGAAAAATGCTTATTTGGGGCAATGCTAAATATTGCGG
CAGTTTTATGCATTGCTACCATTTATGTTTCGTTATAAGCAAGTTCATGCTCTGAGTCCTGAA
GAGAACGTTATCATCAAATTAAACAAGGCTGGCCTTGTAAGTGGAGCTGAGTTGTTTAGG
ACTTTCTATTGTGGCAAACCTCCAGAAAACAACCCTTTTGTCTGCACATGTAAGTGGAGCTG
TGCTTACCTTTGGTATGGGCTCATTATATATGTTTGTTCAGACCATCCTTTCCTACCAAATG
CAGCCCCAAAATCCATGGCAAACAAGTCTTCTGGATCAGACTGTTGTTGGTTATCTGGTGTGG
AGTAAGTGCACCTTAGCATGCTGACTTGCTCATCAGTTTTGCACAGTGGCAATTTTGGGACTG
ATTTAGAACAGAACTCCATTGGAACCCCGAGGACAAAGGTTATGTGCTTCACATGATCACT
ACTGCAGCAGAATGGTCTATGTCATTTTCCTTCTTTGGTTTTTTCCTGACTTACATTCGTGA
TTTTTCAGAAAATTTCTTTACGGGTGGAAGCCAATTTACATGGATTAACCCTCTATGACACTG
CACCTTGCCCTATTAACAATGAACGAACACGGCTACTTTCCAGAGATATTTTGAATGAAAGGAT
AAAATATTTCTGTAATGATTATGATTCTCAGGGATTGGGGAAAGGTTACAGAAGTTGCTTA
TTCTTCTCTGAAATTTTCAACCACTTAATCAAGGCTGACAGTAACACTGATGAATGCTGATA
ATCAGGAAACATGAAAGAAGCCATTTGATAGATTATTCTAAAGGATATCATCAAGAAGACTA
TTAAAAACACCTATGCCTATACTTTTTTATCTCAGAAAATAAAGTCAAAAGACTATG

15/330

FIGURE 15

MWWFQQGLSFLPSALVIWTSAAFI¹FSYITAVTLHHIDPALPYISDTGTVAPEKCLFGAMLNI
AAVLCIATIIYVRYKQVHALSPEENVIIKLNKAGLVLGILSCLGLSIVANFQKTTLFAAHVSG
AVLTFGMGSLYMFVQTILSYQMOPKIHGQVFWIRLLLVIWCGVSALSMLTCSSVLHSGNFG
TDLEQKLHWNPEDKGYVLHMITTAAEWSMSFSFFGFFLT²YIRDFQKISLRVEANLHGLTLYD
TAPCPINNERTRLLSRDI

15/330

FIGURE 16

CGGACGCTTGGGCNCGGCCAGCGGCCAGCGCTAGTCGGTCTGGTAAGTGCCTGATGCCGAGT
TCCGTCTCTCGGGTCTTTTCCTGGTCCCAGGCAAAGCGGAGCGGAGATCCTCAAACGGCCTA
GTGCTTCGCGCTTCCGGAGAAAATCAGCGGTCTAATTAATTCCTCTGGTTTGTTGAAGCAGT
TACCAAGAATCTTCAACCCTTTCCCACAAAAGCTAATTGAGTACACGTTCCCTGTTGAGTACA
CGTTCCTGTTGATTTACAAAAGGTGCAGGTATGAGCAGGTCTGAAGACTAACATTTTGTGAA
GTTGTAAAACAGAAAACCTGTTAGAAATGTGGTGGTTTCAGCAAGGCCTCAGTTTCCTTCCT
TCAGCCCTTGTAATTTGGACATCTGCTGCTTTCATATTTTCATACATTACTGCAGTAACACT
CCACCATATAGACCCGGCTTTACCTTATATCAGTGACACTGGTACAGTANC

FIGURE 17

CCCACGCGTCCGCCCCGCCGTGCGTCCCGGAGTGCAAGTGAGCTTCTCGGCTGCCCCGCGGG
 CCGGGGTGCGGAGCCGAC**ATG**CGCCCGCTTCTCGGCCTCCTTCTGGTCTTCGCCGGCTGCAC
 CTTCGCCTTGTACTTGCTGTGACGCGACTGCCCCGCGGGCGGAGACTGGGCTCCACCGAGG
 AGGCTGGAGGCAGGTCGCTGTGGTTCCCCTCCGACCTGGCAGAGCTGCGGGAGCTCTCTGAG
 GTCCTTCGAGAGTACCGGAAGGAGCACCAGGCCTACGTGTTCTTGCTCTTCTGCGGCGCCTA
 CCTCTACAAACAGGGCTTTGCCATCCCCGGCTCCAGCTTCCTGAATGTTTTAGCTGGTGCCT
 TGTTTGGGCCATGGCTGGGGCTTCTGCTGTGCTGTGTGTTGACCTCGGTGGGTGCCACATGC
 TGCTACCTGCTCTCCAGTATTTTTGGCAAACAGTTGGTGGTGCCTACTTTCTGATAAAGT
 GGCCCTGCTGCAGAGAAAGGTGGAGGAGAACAGAAACAGCTTGTTTTTTTTCTTATTGTTTT
 TGAGACTTTTCCCCATGACACCAAACCTGGTTCTTGAACCTCTCGGCCCCAATTCTGAACATT
 CCCATCGTGCAGTTCTTCTTCTCAGTTCTTATCGGTTTGATCCCATATAATTTTCATCTGTGT
 GCAGACAGGGTCCATCCTGTCAACCCTAACCTCTCTGGATGCTCTTTTCTCCTGGGACACTG
 TCTTTAAGCTGTTGGCCATTGCCATGGTGGCATTAAATTCCTGGAACCCTCATTAAAAAATTT
 AGTCAGAAACATCTGCAATTGAATGAAACAAGTACTGCTAATCATATACACAGTAGAAAAGA
 CACA**TGA**TCTGGATTTTCTGTTTGCCACATCCCTGGACTCAGTTGCTTATTTGTGTAATGGA
 TGTGGTCCTCTAAAGCCCCTCATTGTTTTTGATTGCCTTCTATAGGTGATGTGGACACTGTG
 CATCAATGTGCAGTGTCTTTTCAGAAAGGACACTCTGCTCTTGAAGGTGTATTACATCAGGT
 TTTCAAACCAGCCCTGGTGTAGCAGACACTGCAACAGATGCCTCCTAGAAAATGCTGTTTGT
 GGCCGGGCGCGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCCGGTGATTC
 ACAAGGTCAGGAGTTCAAGACCAGCCTGGCCAAGATGGTGAAATCCTGTCTCTAATAAAAAT
 ACAAAAATTAGCCAGGCGTGGTGGCAGGCACCTGTAATCCCAGCTACTCGGGAGGCTGAGGC
 AGGAGAATTGCTTGAACCAAGGTGGCAGAGGTTGCAGTAAGCCAAGATCACACCACTGCACT
 CCAGCCTGGGTGATAGAGTGAGACACTGTCTTGAC

1000
 900
 800
 700
 600
 500
 400
 300
 200
 100
 0

FIGURE 18

MRPLLGLLLVFAGCTFALYLLSTRLPGRRLGSTEEAGGRSLWFPSDLAELRESEVLREYR
KEHQAYVFLLFCGAYLYKQGFAIPGSSFLNVLGALFGPWLGLLLCCVLTSVGATCCYLLSS
IFGKQLVVSYPDKVALLQRKVEENRNSLFFFLLFLRLFPMTPNWFLNLSAPILNIPVQFF
FSVLIGLIPYNFICVQTGSILSTLTSLDALFSWDTVFKLLAIAMVALIPGTLIKKFSQKHLQ
LNETSTANHIHSRKDT

Important features:

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 101-123, 189-211

N-glycosylation sites.

amino acids 172-176, 250-254

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 240-244, 261-265

N-myristoylation site.

amino acids 13-19, 104-110, 115-121, 204-210

Amidation site.

amino acids 27-31

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 4-15

Protein splicing proteins.

amino acids 25-31

Sugar transport proteins.

amino acids 162-172

18/330

FIGURE 19

CCGAGGCGGGAGGAGCCCGAGGGGGCGCGAGCCCCGCATGAATCATTGTAGTCAATCATTTT
CCAGTTCTCAGCCGCTCAGTTGTGATCAAGGGACACGTGGTTTCCGAAC TGCCAGCTCAGAA
TAGGAAAATAACTTGGGATTTTATATTGGAAGAC **ATG**GATCTTGCTGCCAACGAGATCAGCA
TTTATGACAAACTTTT CAGAGACTGTTGATTTGGTGAGACAGACCGGCCATCAGTGTGGCATG
TCAGAGAAGGCAATTGAAAAATTTATCAGACAGCTGCTGGAAAAGAATGAACCTCAGAGACC
CCCCCGCAGTATCCTCTCCTTATAGTTGTGTATAAGGTTCTCGCAACCTTGGGATTAATCT
TGCTCACTGCCTACTTTGTGATTTCAACCTTT CAGCCCATTAGCACCTGAGCCAGTGCTTTCT
GGAGCTCACACCTGGCGCTCACTCATCCATCACATTAGGCTGATGTCCTTGCCCATTGCCAA
GAAGTACATGTCAGAAAATAAGGGAGTTCCTCTGCATGGGGGTGATGAAGACAGACCCTTTC
CAGACTTTGACCCCTGGTGGACAAACGACTGTGAGCAGAATGAGTCAGAGCCCATTCTGCC
AACTGCACTGGCTGTGCCCAGAAACACCTGAAGGTGATGCTCCTGGAAGACGCCCCAAGGAA
ATTTGAGAGGCTCCATCCACTGGTGATCAAGACGGGAAAGCCCCTGTTGGAGGAAGAGATTCT
AGCATTTTTTGTGCCAGTACCCTGAGGCGACAGAAGGCTTCTCTGAAGGGTTTTTCGCCAAG
TGGTGGCGCTGCTTTCCTGAGCGGTGGTTCCCATTTCTTATCCATGGAGGAGACCTCTGAA
CAGATCACAAATGTTACGTGAGCTTTTTCTGTTTTCACTCACCTGCCATTTCCAAAAGATG
CCTCTTTAAACAAGTGCTCCTTTCTTCACCCAGAACCTGTTGTGGGGAGTAAGATGCATAAG
ATGCCTGACCTATTTATCATTGGCAGCGGTGAGGCCATGTTGCAGCTCATCCCTCCCTTCCA
GTGCCGAAGACATTGTCAGTCTGTGGCCATGCCAATAGAGCCAGGGGATATCGGCTATGTCTG
ACACCACCCACTGGAAGGTCTACGTTATAGCCAGAGGGGTCCAGCCTTTGGTCATCTGCGAT
GGAACCGCTTTCTCAGAACTG **TAG**GAAATAGAACTGTGCACAGGAACAGCTTCCAGAGCCGA
AAACCAGGTTGAAAGGGGAAAAATAAAAACAAAAACGATGAAACTGCAAAAA

FIGURE 20

MDLAANEISIIYDKLSETVDLVRQTGHQCGMSEKAIEKFIRQLLEKNEPQRPPPQYPLLIVVY
KVLATLGLILLTAYFVIQPFSPLAPEPVLSGAHTWRSLIHHIRLMSLPIAKKYMSENKGVPL
HGGDEDRPFPDFDPWWTNDCEQNESEPIPANCTGCAQKHLKVMLLEDAPRKFERLHPLVIKT
GKPLLEEEIQHFLCQYPEATEGFSEGGFAKWWRCFPERWFPPYPWRRPLNRSQMLRELPV
FTHLPFPKDASLNKCSFLHPEPVVGSKMHKMPDLFIIGSGEAMLQLIPPFQCRHCQSVAMP
IEPGDIGYVDTTTHWKVYVIARGVQPLVICDGTAFSEL

20/330

FIGURE 21

CCACGGTGTCCGTTCTTCGCCCCGGCGGCAGCTGTCCCCGAGGCGGGAGGAGCCCCGAGGGGCG
CGAGCCCCGCATGAATCATTGTAGTCAATCATTTTCCAGTTCTCAGCCGTTTCAGTTGTGATC
AAGGGACACGTGGTTTCCGAACTGCCAGCTCAGAATAGGAAAATAACTTGGGATTTTATATT
GGAAGACATGGATCTTGCTGCCAACGAGATCAGCATTTATGACAACTTTCAGAGACTGTTG
ATTTGGTGAGACAGACCGGCCATCAGTGTGGCATGTGAGAGAAGGCAATTGAAAAATTTATC
AGACAGCTGCTGGAAAAGAAATGAACCTCAGAGACCCCCCCCCGAGTATCCTCTCCTTATAGT
TGTGTATAAGGTTCTCGCAACCTTGGGATTAATCTTGCTCACTGCCTACTTTGTGATTCAAC
CTTTCAGCCCATTAGCACCTGAGCCAGTGCTTTGTGGAGCTCAC

CCACGGTGTCCGTTCTTCGCCCCGGCGGCAGCTGTCCCCGAGGCGGGAGGAGCCCCGAGGGGCG
CGAGCCCCGCATGAATCATTGTAGTCAATCATTTTCCAGTTCTCAGCCGTTTCAGTTGTGATC
AAGGGACACGTGGTTTCCGAACTGCCAGCTCAGAATAGGAAAATAACTTGGGATTTTATATT
GGAAGACATGGATCTTGCTGCCAACGAGATCAGCATTTATGACAACTTTCAGAGACTGTTG
ATTTGGTGAGACAGACCGGCCATCAGTGTGGCATGTGAGAGAAGGCAATTGAAAAATTTATC
AGACAGCTGCTGGAAAAGAAATGAACCTCAGAGACCCCCCCCCGAGTATCCTCTCCTTATAGT
TGTGTATAAGGTTCTCGCAACCTTGGGATTAATCTTGCTCACTGCCTACTTTGTGATTCAAC
CTTTCAGCCCATTAGCACCTGAGCCAGTGCTTTGTGGAGCTCAC

FIGURE 22

CCCACGCGTCCGCCCCACGCGTCCGGCTGAACACCTCTTCTTTGGAGTCAGCCACTGATGAGG
 CAGGGTCCCCACTTGCAGCTGCAGCAGCTGCAGCAGCTGCAGAGCGCTGCTCCTGGCTGGTG
 CCACTGGTGCGCACGCTGCTAGACCGTGCCTATGAGCCGCTGGGGCTGCAGTGGGGACTGCC
 CTCCCTGCCACCCACCAATGGCAGCCCCACCTTCTTTGAAGACTTCCAGGCTTTTTGTGCCA
 CACCCGAATGGCGCCACTTTCATCGACAAACAGGTACAGCCAACC**ATG**TCCAGTTCGAAATG
 GACACGTATGCTAAGAGCCACGACCTTATGTGAGGTTTCTGGAATGCCTGCTATGACATGCT
 TATGAGCAGTGGGCAGCGGCCAGTGGGAGCGCGCCAGAGTCGTGCGGCCCTTCCAGGAGC
 TGGTGCTGGAACCTGCGCAGAGGCGGGCGCGCTGGAGGGGCTACGCTACACGGCAGTGCTG
 AAGCAGCAGGCAACGCAGCACTCCATGGCCCTGCTGCACTGGGGGGCGCTGTGGCGCCAGCT
 CGCCAGCCCATGTGGGGCCTGGGCGCTGAGGGACACTCCCATCCCCCGCTGGAAACTGTCCA
 GCGCCGAGACATATTACGCGATGCGTCTGAAGCTGGTGCCCAACCATCACTTCGACCCCTCAC
 CTGGAAGCCAGCGCTCTCCGAGACAACTCTGGGTGAGGTTCCCTGACACCCACCGAGGAGGC
 CTCACCTGCCTCTGGCAGTGACCAAAGAGGGCCAAAGTGAGCACCCACCCGAGTTGCTGCAGG
 AGGACCAGCTCGGCGAGGACGAGCTGGCTGAGCTGGAGACCCGATGGAGGCAGCAGAACTG
 GATGAGCAGCGTGAGAAGCTGGTGCTGTCGGCCGAGTGCCAGCTGGTGACGGTAGTGGCCGT
 GGTCCCAGGGCTGCTGGAGGTCACCACAGAAATGTATACTTCTACGATGGCAGCACTGAGC
 GCGTGGAACCCGAGGAGGGCATCGGCTATGATTTCCGGCGCCCACTGGCCAGCTGCGTGAG
 GTCCACCTGCGGCGTTTCAACCTGCGCGCTTACGCACTTGAGCTCTTCTTTATCGATCAGGC
 CAACTACTTCTCAACTTCCCATGCAAGGTGGGCACGACCCCACTCTCATCTCCTAGCCAGA
 CTCCGAGACCCAGCCTGGCCCCATCCCACCCCATACCCAGGTACGGAACCAAGGTGTACTCG
 TGGCTCCTGCGCCTACGGCCCCCTCTCAAGGTACCTAAGCAGCCGCTCCCCCAGGAGAT
 GCTGCGTGCTCAGGCCCTTACCCAGAAATGGGTACAGCGTGAGATATCCAACCTTCGAGTACT
 TGATGCAACTCAACACCATTTGCGGGGCGGACCTACAATGACCTGTCTCAGTACCCTGTGTTC
 CCCTGGGTCTGTCAGGACTACGTGTCCCCAACCCTGGACCTCAGCAACCCAGCCGTCTTCCG
 GGACCTGTCTAAGCCCATCGGTGTGGTGAACCCCAAGCATGCCAGCTCGTGAGGGAGAAGT
 ATGAAAGCTTTGAGGACCCAGCAGGGACCATTGACAAGTTCCACTATGGCACCCCACTACTCC
 AATGCAGCAGGCGTGATGCCTACCTCATCCGCGTGGAGCCCTTCACTCCCTGAGCCTCCA
 GCTGCAAAGTGGCCGCTTTGACTGCTCCGACCGGCAGTTCCACTCGGTGGCGGCAGCCTGGC
 AGGCACGCCTGGAGAGCCCTGCCGATGTGAAGGAGCTCATCCCGGAATTTCTTCTACTTTCT
 GACTTCCTGGAGAACCAGAACGGTTTTGACCTGGGCTGTCTCCAGCTGACCAACGAGAAGGT
 AGGCGATGTGGTGCTACCCCCGTGGGCCAGCTCTCCTGAGGACTTCATCCAGCAGCCAGCC
 AGGCTCTGGAGTCGGAGTATGTCTGCACACCTACACGAGTGGATCGACCTCATCTTTGGC
 TACAAGCAGCGGGGGCCAGCGCCGAGGAGGCCCTCAATGTCTTCTATTACTGCACCTATGA
 GGGGGCTGTAGACCTGGACCATGTGACAGATGAGCGGGAACGGAAGGCTCTGGAGGGCATTA
 TCAGCAACTTTGGGCAGACTCCCTGTGAGCTGCTGAAGGAGCCACATCCAACCTCGGCTCTCA
 GCTGAGGAAGCAGCCCATCGCCTTGACGCCTGGACACTAACTCACCTAGCATCTTCCAGCA
 CCTGGACGAATCAAGGCATTCTTCGAGAGGGTGACTGTGAGTGCCAGTGGGCTGCTGGGCA
 CCCACAGCTGGTTGCCCTATGACCGCAACATAAGCAACTACTTCAGCTTCAGCAAAGACCCC
 ACCATGGGCAGCCACAAGACGCAGCGACTGCTGAGTGGCCCGTGGGTGCCAGGCAGTGGTGT
 GAGTGGACAAGCACTGGCAGTGGCCCCGGATGGAAAGCTGCTATTACGCGGTGGCCACTGGG
 ATGGCAGCCTGCGGGTGACTGCACTACCCCGTGGCAAGCTGTTGAGCCAGCTCAGCTGCCAC
 CTTGATGTAGTAACCTGCCTTGCACTGGACACCTGTGGCATCTACCTCATCTCAGGCTCCCG
 GGACACCACGTGCATGGTGTGGCGGCTCCTGCATCAGGGTGGTCTGTGAGTAGGCCTGGCAC
 CAAAGCCTGTGCAGGTCTGTATGGGCATGGGGCTGCAGTGAGCTGTGTGGCCATCAGCACT
 GAACTTGACATGGCTGTGTCTGGATCTGAGGATGGAAGTGTGATCATAACACTGTACGCCG
 CGGACAGTTTGTAGCGGCACTACGGCCTCTGGGTGCCACATTCCTGGACCTATTTTCCACC
 TGGCATTGGGGTCCGAAGGCCAGATTGTGGTACAGAGCTCAGCGTGGAACGTCTCTGGGGCC
 CAGGTCACCTACTCCTTGACCTGTATTGAGTCAATGGGAAGTTGCGGGCTTCACTGCCCT
 GGCAGAGCAGCCTACAGCCCTGACGGTGACAGAGGACTTTGTGTTGCTGGGCACCGCCAGT
 GCGCCCTGCACATCCTCCAACATAACACTGCTCCCGGCCGCGCCTCCCTTGCCCATGAAG
 GTGGCCATCCGCAGCGTGGCCGTGACCAAGGAGCGCAGCCACGTGCTGGTGGGCTGGAGGA
 TGGCAAGCTCATCGTGGTGGTGGCGGGGACCCCTCTGAGGTGCGCAGCAGCCAGTTCGCGC
 GGAAGCTGTGGCGTCTCCTCGCGGCGCATCTCCAGGTGTCTCGGGAGAGACGGAATACAAC
 CCTACTGAGGCGCGCT**TGA**ACCTGGCCAGTCCGGCTGCTCGGGCCCCGCCCGGCGAGGCCTG
 GCCCGGGAGGCCCCGCCAGAGTCGGCGGGAACACCCCGGGGTGGGCAGCCAGGGGGTGA
 GCGGGGGCCACCTGCCAGCTCAGGGATTGGCGGGCGATGTTACCCCTCAGGGATTGGCG
 GGCGGAAGTCCCGCCCCCTCGCCGGCTGAGGGGGCCGCCCTGAGGGGCCAGCACTGGCGTCT

FIGURE 23

MSQFEMDTYAKSHDLMSGFWNACYDMLMSSGQRRQWERAQSRRAFQELVLEPAQRRARLEGL
 RYTAVLKQQATQHSMALLHWGALWRQLASPCGAWALRDTPIPRWKLSSAETYSRMRLKLVN
 HHFDPHLEASALRDNLGEVPLTPTEEASLPLAVTKEAKVSTPPELLQEDQLGEDELALETP
 MEAAELDEQREKLVLSAECQLVTVVAVVPGLLEVTTQNVYFYDGSTERVETEEGIGYDFRRP
 LAQLREVHLRRFNLRRSALELEFFIDQANYFLNFPCKVGTTPVSSPSQTPRPQPGPIPPHTQV
 RNQVYSWLLRLRPPSQGYLSSRSPQEMLRASGLTQKWVQREISNFEYLMQLNTIAGRTYNDL
 SQYPVFPWVLQDYVSPTLDLSNPAVFRDLSKPIGVVNPKHAQLVREKYESFEDPAGTIDKFH
 YGTHYSNAAGVMHYLIRVEPFTSLHVQLQSGRFDCSDRQFHSVAAAWQARLESPADVKEIP
 EFFYFPDFLENQNGFDLGCLQLTNEKVGDVVLPPWASSPEDFIQQHRQALESEYVSAHLHEW
 IDLIFGYKQRGPAEEALNVFYCYEGAVDLDHVTDERERKALEGIISNFGQTPCQLLKEP
 HPTRLSAEEAAHRLARLDTNSPSIFQHLDELKAFFAEVTVSASGLLGTHSWLPYDRNISNYF
 SFSKDPTMGSHKTQRLLSGPWVPGSGVSGQALAVAPDGKLLFSGGHWGDSLRTALPRGKLL
 SQLSCHLDVVTCLALDTCGIYLISGSRDTTCMVWRLHQQGLSVGLAPKPVQVLYGHGAAVS
 CVAISTELDMAVSGSEDGTVIIHTVRRGQFVAALRPLGATFPGPFIHFLALGSEGQIVVQSSA
 WERPGAQVTYSLHLYSVNGKLRLASLPLAEQPTALTVTEDFVLLGTAQCALHILQLNTLLPAA
 PPLPMKVAIRSVAVTKERSHVLVGLLEDGKLIVVAVAGQPSEVRSSQFARKLWRSSRRISQVSS
 GETEYNPTAR

N-glycosylation site.

amino acids 677-681

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 985-989

Tyrosine kinase phosphorylation site.

amino acids 56-65, 367-376, 543-551

N-myristoylation site.

amino acids 61-67, 436-442, 604-610, 610-616, 664-670, 691-697,
 706-712, 711-717, 769-775, 785-791, 802-808, 820-826, 834-840,
 873-879, 912-918, 954-960

FIGURE 24

CGGACGCGTGGGCGGACGCGTGGGGGCTGTGAGAAAGTGCCAATAAATACATCATGCAACCC
CACGGCCCACCTTGTGAACTCCTCGTGCCCAGGGCTGATGTGCGTCTTCCAGGGCTACTCAT
CCAAAGGCCTAATCCAACGTTCTGTCTTCAATCTGCAAATCTATGGGGTCCTGGGGCTCTTC
TGGACCCCTTAAGTGGTACTGGCCCTGGGCCAATGCGTCCTCGCTGGAGCCTTGCCTCCTT
CTACTGGGCCTTCCACAAGCCCCAGGACATCCCTACCTTCCCCTTAATCTCTGCCTTCATCC
GCACACTCCGTTACCACACTGGGTCAATTGGCATTTGGAGCCCTCATCCTGACCCTTGTGCAG
ATAGCCCGGGTCATCTTGGAGTATATTGACCACAAGCTCAGAGGAGTGCAGAACCCTGTAGC
CCGCTGCATCATGTGCTGTTTCAAGTGCTGCCTCTGGTGTCTGGAAAAATTTATCAAGTTCC
TAAACCGCAATGCATACATCATGATCGCCATCTACGGGAAGAATTTCTGTGTCTCAGCCAAA
AATGCGTTCATGCTACTCATGCGAAACATTGTCAGGGTGGTCGTCCTGGACAAAGTCACAGA
CCTGCTGCTGTTCTTTGGGAAGCTGCTGGTGGTCGGAGGCGTGGGGGTCCTGTCCTTCTTTT
TTTTCTCCGGTCGCATCCCGGGGCTGGGTAAAGACTTTAAGAGCCCCCACCTCAACTATTAC
TGGCTGCCCATCATGACCTCCATCCTGGGGGCCTATGTCATCGCCAGCGGCTTCTTCAGCGT
TTTCGGCATGTGTGTGGACACGCTCTTCCTCTGCTTCTGGAAGACCTGGAGCGGAACAACG
GCTCCCTGGACCGGCCCTACTACATGTCCAAGAGCCTTCTAAAGATTCTGGGCAAGAAGAAC
GAGGCGCCCCCGGACAACAAGAAGAGGAAGAAGTGACAGCTCCGGCCCTGATCCAGGACTGC
ACCCACCCCCACCGTCCAGCCATCCAACCTCACTTCGCCTTACAGGTCTCCATTTTGTGGT
AAAAAAGGTTTTAGGCCAGGCGCCGTGGCTCACGCCTGTAATCCAACACTTTGAGAGGCTG
AGGCGGGCGGATCACCTGAGTCAGGAGTTCGAGACCAGCCTGGCCAACATGGTGAAACCTCC
GTCTCTATTAAAAATACAAAAATTAGCCGAGAGTGGTGGCATGCACCTGTCATCCCAGCTAC
TCGGGAGGCTGAGGCAGGAGAATCGCTTGAACCCGGGAGGCAGAGGTTGCAGTGAGCCGAGA
TCGCGCCACTGCACTCCAACCTGGGTGACAGACTCTGTCTCCAAAACAAAACAAACAA
AAAGATTTTATTAAAGATATTTTGTAACTC

FIGURE 25

RTRGRTRGGCEKVPINTSCNPTAHLVNSSCPGLMCVFQGYSSKGLIQRSVFNLQIYGVLGLF
WTLNWWLALGQCVLAGAFASFYWAFHKPQDIPTFPLISAFIRTLRYHTGSLAFGALILTLVQ
IARVILEYIDHKLRGVQNPVARCIMCCFKCCLWCLEKFIKFLNRNAYIMIAIYGKNFCVSAK
NAFMLLMRNIVRVVLDKVTDLLFFGKLLVVGVGVLSTFFFSGRIPGLGKDFKSPHLNYY
WLPIMTSILGAYVIASGFFSVFGMCVDTLFLCFLEDLERNNGSLDRPYMSKSLLKILGKKN
EAPPDNKKRKK

25/330

FIGURE 26

GAGTCTTGACCGCCGCCGGGCTCTTGGTACCTCAGCGCGAGCGCCAGGCGTCCGGCCGCCGT
 GGCTATGTTCGTGTCCGATTTCCGCAAAGAGTTCTACGAGGTGGTCCAGAGCCAGAGGGTCC
 TTCTCTTCGTGGCCTCGGACGTGGATGCTCTGTGTGCGTGCAAGATCCTTCAGGCCTTGTTT
 CAGTGTGACCACGTGCAATATACGCTGGTTCCAGTTTCTGGGTGGCAAGAACTTGAACTGC
 ATTTCTTGAGCATAAAGAACAGTTTCATTATTTTATTCTCATAAACTGTGGAGCTAATGTAG
 ACCTATTGGATATTCTTCAACCTGATGAAGACACTATATTCTTTGTGTGTGACTCCCATAGG
 CCAGTCAATGTCGTCAATGTATACAACGATACCCAGATCAAATTAATCATTAAACAAGATGA
 TGACCTTGAAGTTCCCGCCTATGAAGACATCTTCAGGGATGAAGAGGAGGATGAAGAGCATT
 CAGGAAATGACAGTGATGGGTGAGAGCCTTCTGAGAAGCGCACACGGTTAGAAGAGGAGATA
 GTGGAGCAAACCATGCGGAGGAGGCAGCGGCGAGAGTGGGAGGCCCGGAGAAGAGACATCCT
 CTTTGACTACGAGCAGTATGAATATCATGGGACATCGTCAGCCATGGTGATGTTTGAGCTGG
 CTTGGATGCTGTCCAAGGACCTGAATGACATGCTGTGGTGGGCCATCGTTGGACTAACAGAC
 CAGTGGGTGCAAGACAAGATCACTCAAATGAAATACGTGACTGATGTTGGTGTCTTGCAGCG
 CCACGTTTCCCGCCACAACCACCGGAACGAGGATGAGGAGAACACACTCTCCGTGGACTGCA
 CACGGATCTCCTTTGAGTATGACCTCCGCCTGGTGCTCTACCAGCACTGGTCCCTCCATGAC
 AGCCTGTGCAACACCAGCTATACCGCAGCCAGGTTCAAGCTGTGGTCTGTGCATGGACAGAA
 GCGGCTCCAGGAGTTCCTTGACAGCATGGGTCTTCCCCTGAAGCAGGTGAAGCAGAAGTTCC
 AGGCCATGGACATCTCCTTGAAAGGAGAATTTGCGGGAAATGATTGAAGAGTCTGCAAATAAA
 TTTGGGATGAAGGACATGCGCGTGCAGACTTTCAGCATTCAATTTGGGTTCAGCACAAAGTT
 TCTGGCCAGCGACGTGGTCTTTGCCACCATGTCTTTGATGGAGAGCCCCGAGAAGGATGGCT
 CAGGGACAGATCACTTCATCCAGGCTCTGGACAGCCTCTCCAGGAGTAACCTGGACAAGCTG
 TACCATGGCCTGGAACCTGCCAAGAAGCAGCTGCGAGCCACCCAGCAGACCATTGCCAGCTGC
 CTTTGACCAACCTCGTCATCTCCCAGGGGCCTTTCCTGTACTGCTCTCTCATGGAGGGCAC
 TCCAGATGTCATGCTGTTCTCTAGGCCGGCATCCCTAAGCCTGCTCAGCAAACACCTGCTCA
 AGTCCTTTGTGTGTTGACAAAGAACCGGCGCTGCAAACCTGCTGCCCCTGGTGATGGCTGCC
 CCCCTGAGCATGGAGCATGGCACAGTGACCGTGGTGGGCATCCCCCAGAGACCGACAGCTC
 GGACAGGAAGAACTTTTTTGGGAGGGCGTTTGAGAAGGCAGCGGAAAGCACCAGCTCCCGGA
 TGCTGCACAACCATTTTGACCTCTCAGTAATTGAGCTGAAAGCTGAGGATCGGAGCAAGTTT
 CTGGACGCACTTATTTCCCTCCTGTCCTAGGAATTTGATTCTTCCAGAATGACCTTCTTATT
 TATGTAACCTGGCTTTCATTTAGATTGTAAGTTATGGACATGATTTGAGATGTAGAAGCCATT
 TTTTATTAAATAAAATGCTTATTTTAGGAAA

27/330

FIGURE 27

MFVSDERKEFYEVVQSQRVLLFVASDVDALCACKILQALFQCDHVQYTLVPVSGWQELETAF
LEHKEQFHYFILINCGANVDLLDILQPDEDITFFVCDSHRPVNVNVYNDTQIKLLIKQDDD
LEVPAIEDIFRDEEEDDEHSGNDSGDSEPSEKRTRLEEEIVEQTMRRRQRREWEARRRDILF
DYEQYEHGTSSAMVMFELAWMLSKDLNDMLWWAIVGLTDQWVQDKITQMKYVTDVGVLRH
VSRHNRNEDEENTLSVDCTRISFEYDLRLVLYQHWSLHDSL CNTSYTAARFKLWSVHGQKR
LQEF LADMGLPLKQVKQKFQAMD ISLKENLREMIEESANKFGMKDMRVQTF SIHFGFKHKFL
ASDVVFATMSLMESPEKDGSGTDHFIQALDSLRSNLDKLYHGLELAKKQLRATQQTIASCL
CTNLVISQGPFLYCSLMEGTPDVMLFSRPASLSLLSKHLLKS FVCSTKNRRCKLLPLVMAAP
LSMEHGTVTVVGIPPETDSSDRKNFFGRAFEKAAESTSSRMLHNNHFDLSVIELKAEDRSKFL
DALISLLS

FIGURE 28

GTACCTCAGCGCGAGCGCCAGGCGTCCGGCCGCCGTGGCTATGNTCGTGTCCGATTTCCGCA
AAGAGTTCTACGAGGTGGTCCAGAGCCAGAGGGTCCTTCTCTTCGTGGCCTCGGANGTGGAT
GCTCTGTGTGCGTGCAAGATCCTTCAGGCCTTGTTCCAGTGTGACCANGTGCAATATANGCT
GGTTCAGTTTCTGGGTGGCAAGAACTTGAAACTGCATTTCTTGAGCATAAAGAACAGTTTC
ATTATTTTATTCTCATAAACTGTGGAGCTAATGTAGACCTATTGGATATTCTTCAACCTGAT
GAAGACACTATATTCTTTGTGTGTGACACCCATAGGCCAGTCAATGTTGTCAATGTATACAA
CGATACCC

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 29

FIGURE 30

MVTAALGPVWAALLLFLLMCEIRMVELTFDRAVASGCQRCCDSEDPLDPAHVSSASSSSGRPH
ALPEIRPYINITILKGDKGDPGPMGLPGYMGREGPQGEPGPQGSKGDKGEMGSPGAPCQKRF
FAFSVGRKTALHSGEDFQTLLFERVFVNLDGCFDMATGQFAAPLRGIYFFSLNVHSWNYKET
YVHIMHNQKEAVILYAQPSESRIMQSQSVMLDLAYGDRVWVRLFKRQRENAIYSNDFDTYIT
FSGHLIKAEDD

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 72-75

C1q domain proteins.

amino acids 144-178, 78-111 and 84-117

ACTCGAACGCAGTTGCTTCGGGACCCAGGACCCCTCGGGCCCCGACCCGCCAGGAAAGACTG
AGGCCGCGGCCTGCCCCGCCCGGCTCCCTGCGCCGCCCGCCGCTCCCGGGACAGAAGATGTG
CTCCAGGGTCCCTCTGCTGCTGCCGCTGCTCCTGCTACTGGCCCTGGGGCCTGGGGTGCAGG
GCTGCCCATCCGGCTGCCAGTGCAGCCAGCCACAGACAGTCTTCTGCACTGCCCGCCAGGGG
ACCACGGTGCCCCGAGACGTGCCACCCGACACGGTGGGGCTGTACGTCTTTGAGAACGGCAT
CACCATGCTCGACGCAGGCAGCTTTGCCGGCCTGCCGGGCCTGCAGCTCCTGGACCTGTAC
AGAACCAGATCGCCAGCCTGCCCAGCGGGGTCTTCCAGCCACTCGCCAACCTCAGCAACCTG
GACCTGACGGCCAACAGGCTGCATGAAATCACCAATGAGACCTTCCGTGGCCTGCGGCGCCT
CGAGCGCCTCTACCTGGGCAAGAACCGCATCCGCCACATCCAGCCTGGTGCCTTCGACACGC
TCGACCGCCTCCTGGAGCTCAAGCTGCAGGACAACGAGCTGCGGGCACTGCCCCCGCTGCGC
CTGCCCCGCCTGCTGCTGCTGGACCTCAGCCACAACAGCCTCCTGGCCCTGGAGCCCGGCAT
CCTGGACACTGCCAACGTGGAGGCGCTGCGGCTGGCTGGTCTGGGGCTGCAGCAGCTGGACG
AGGGGCTCTTTCAGCCGCTTGCGCAACCTCCACGACCTGGATGTGTCCGACAACCAGCTGGAG
CGAGTGCCACCTGTGATCCGAGGCCTCCGGGGCCTGACGCGCCTGCGGCTGGCCGGCAACAC
CCGCATTGCCCAGCTGCGGCCCGAGGACCTGGCCGGCCTGGCTGCCCTGCAGGAGCTGGATG
TGAGCAACCTAAGCCTGCAGGCCCTGCCTGGCGACCTCTCGGGCCTCTTCCCCGCCTGCGG
CTGCTGGCAGCTGCCCCGAACCCCTTCAACTGCGTGTGCCCCCTGAGCTGGTTTGGCCCTG
GGTGC GCGAGAGCCACGTCACTGGCCAGCCCTGAGGAGACGCGCTGCCACTTCCCGCCCA
AGAACGCTGGCCGGCTGCTCCTGGAGCTTGACTACGCCGACTTTGGCTGCCCAGCCACCACC
ACCACAGCCACAGTGCCCAACCACGAGGCCCGTGGTGC GGGAGGCCACAGCCTTGTCTTCTAG
CTTGGCTCCTACCTGGCTTAGCCCCACAGCGCCGGCCACTGAGGCCCCAGCCGCCCCCTCCA
CTGCCCCACCGACTGTAGGGCCTGTCCCCCAGCCCCAGGACTGCCCACCCTCCACCTGCCTC
AATGGGGGACATGCCACCTGGGGACACGGCACCACCTGGCGTGCTTGTGCCCCGAAGGCTT
CACGGGCCTGTACTGTGAGAGCCAGATGGGGCAGGGGACACGGCCCAGCCCTACACCAGTCA
CGCCGAGGCCACCACGGTCCCTGACCCTGGGCATCGAGCCGGTGAGCCCCACCTCCCTGCGC
GTGGGGCTGCAGCGCTACCTCCAGGGGAGCTCCGTGCAGCTCAGGAGCCTCCGTCTCACCTA
TCGCAACCTATCGGGCCCTGATAAGCGGCTGGTGACGCTGCGACTGCCTGCCTCGCTCGCTG
AGTACACGGTCACCCAGCTGCGGCCCAACGCCACTTACTCCGTCTGTGTATGCCTTTGGGG
CCCCGGGCGGGTGCCGGAGGGCGAGGAGGCCTGCGGGGAGGCCCATACCCCCAGCCGTCCA
CTCCAACCACGCCCCAGTCACCCAGGCCCGCGAGGGCAACCTGCCGCTCCTCATTGCGCCCC
CCCTGGCCGCGGTGCTCCTGGCCGCGCTGGCTGCGGTGGGGGCAGCCTACTGTGTGCGGCGG
GGGCGGGCCATGGCAGCAGCGGCTCAGGACAAAGGGCAGGTGGGGCCAGGGGCTGGGCCCCCT
GGAAGTGGAGGGAGTGAAGGTCCCCTTGAGGCCAGGCCCGAAGGCAACAGAGGGCGGTGGAG
AGGCCCTGCCCAGCGGTCTGAGTGTGAGGTGCCACTCATGGGCTTCCCAGGGCCTGGCCTC
CAGTCACCCCTCCACGCAAAGCCCTACATCTAAAGCCAGAGAGAGACAGGGCAGCTGGGGCCG
GGCTCTCAGCCAGTGAGATGGCCAGCCCCCTCCTGCTGCCACACCACGTAAGTTCTCAGTCC
CAACCTCGGGGATGTGTGCAGACAGGGCTGTGTGACCACAGCTGGGCCCTGTTCCCTCTGGA
CCTCGGTCTCCTCATCTGTGAGATGCTGTGGCCCAGCTGACGAGCCCTAACGTCCCCAGAAC
CGAGTGCCCTATGAGGACAGTGTCCGCCCTGCCCTCCGCAACGTGCAGTCCCTGGGCACGGCG
GGCCCTGCCATGTGCTGGTAACGCATGCCTGGGTCTGCTGGGCTCTCCCACTCCAGGCGGA
CCCTGGGGGCCAGTGAAGGAAGCTCCCGGAAAGAGCAGAGGGAGAGCGGGTAGGCGGCTGTG
TGACTCTAGTCTTGGCCCCAGGAAGCGAAGGAACAAAAGAACTGGAAAGGAAGATGCTTTA
GGAACATGTTTTGCTTTTTTAAATATATATATTTATAAGAGATCCTTTCCCATTTATTCTG
GGAAGATGTTTTTCAAACCTCAGAGACAAGGACTTTGGTTTTTTGTAAGACAAACGATGATATG
AAGGCCTTTTGTAAAGAAAAAATAAAAGATGAAGTGTGAAA

FIGURE 32

MCSRVPLLLLPLLLLLLALGPGVQGCPSGCQCSQPQTVFCTARQGTTPRDVPPDTVGLYVFEN
GITMLDAGSFAGLPGLQLLDLSQNQIASLPSGVFQPLANLSNLDLTANRLHEITNETFRGLR
RLERLYLGKNRIRHIQPGAFDTLDRILLELKLQDNELRALPPLRLPRLLLLDLSHNSLLALEP
GILDTANVEALRLAGLGLQQLDEGLFSRLRNLHDLDVSDNQLERVPPVIRGLRGLTRLRLAG
NTRIAQLRPEDLAGLAALQEILDVSNLSLQALPGDLSGLFPRLRLLAAARNPFNCVCPLSWFG
PWVRESHVTLASPEETRCHFFPKNAGRLLLELDYADFGCPATTTTATVPTTRPVVREPTALS
SSLAPTWLSPTAPATEAPSPPSTAPPTVGPVPQPQDCPPSTCLNGGTCHLGRHHLACLCPE
GFTGLYCESQMGQGTRPSPTPVTPRPPRSLTLGIEPVSPSTSLRVGLQRYLQGSSVQLRSLRL
TYRNLSGPDKRLVTLRLPASLAEYTVTQLRPNATYSVCMPLGPGRVPEGEEACGEAHTPPA
VHSNHAPVTQAREGNLPLLIAPALAAVLLAALAAGAAAYCVRGRAMAAAAQDKGQVGPAGAG
PLELEGVKVPLEPGPKATEGGGEALPSGSECEVPLMGFPGPGLQSPLHAKPYI

FIGURE 33

GAATCATCCACGCACCTGCAGCTCTGCTGAGAGAGTGCAAGCCGTGGGGGTTTTGAGCTCAT
CTTCATCATTATATGAGGAAATAAGTGGTAAATCCTTGGAAATACA**ATG**AGACTCATCAG
AAACATTTACATATTTTGTAGTATTGTTATGACAGCAGAGGGTGATGCTCCAGAGCTGCCAG
AAGAAAGGGAACTGATGACCAACTGCTCCAACATGTCTCTAAGAAAGGTTCCCGCAGACTTG
ACCCAGCCACAACGACACTGGATTTATCCTATAACCTCCTTTTTCAACTCCAGAGTTCAGA
TTTTCATTTCTGTCTCCAACTGAGAGTTTTGATTCTATGCCATAACAGAATTCAACAGCTGG
ATCTCAAAACCTTTGAATTTCAACAAGGAGTTAAGATATTTAGATTTGTCTAATAACAGACTG
AAGAGTGTAACTTGGTATTTACTGGCAGGTCTCAGGTATTTAGATCTTTCTTTTAATGACTT
TGACACCATGCCTATCTGTGAGGAAGCTGGCAACATGTCACACCTGGAAATCCTAGGTTTGA
GTGGGGCAAAAATACAAAATCAGATTTCCAGAAAATTGCTCATCTGCATCTAAATACTGTC
TTCTTAGGATTTCAGAACTCTTCCTCATTATGAAGAAGGTAGCCTGCCCATCTTAAACACAAC
AAAAGTGCACATTGTTTTACCAATGGACACAAATTTCTGGGTTCTTTTGGTGGATGGAATCA
AGACTTCAAAAATATTAGAAATGACAAATATAGATGGCAAAAGCCAATTTGTAAGTTATGAA
ATGCAACGAAATCTTAGTTTTAGAAAATGCTAAGACATCGGTTCTATTGCTTAATAAAGTTGA
TTTACTCTGGGACGACCTTTTCTTATCTTACAATTTGTTTGGCATAACATCAGTGGAAACACT
TTCAGATCCGAAATGTGACTTTTGGTGGTAAGGCTTATCTTGACCACAATTCATTTGACTAC
TCAAATACTGTAATGAGAACTATAAAATGGAGCATGTACATTTAGAGTGTTTTACATTCA
ACAGGATAAAAATCTATTTGCTTTTGACCAAAATGGACATAGAAAACCTGACAATATCAAATG
CACAAATGCCACACATGCTTTTCCCGAATTATCCTACGAAATCCAATATTTAAATTTTGCC
AATAATATCTTAACAGACGAGTTGTTTAAAAGAACTATCCAACCTGCCTCACTTGAAAACCTCT
CATTTTGAATGGCAATAAACTGGAGACACTTTCTTTAGTAAGTTGCTTTGCTAACAACACAC
CCTTGGAACACTTGGATCTGAGTCAAAATCTATTACAAACATAAAAATGATGAAAATTGCTCA
TGGCCAGAACTGTGGTCAATATGAATCTGTATACAAATAAAATTTGTCTGATTCTGTCTTCAG
GTGCTTGCCCCAAAAGTATTCAAATACTTGACCTAAATAATAACCAAATCCAAACTGTACCTA
AAGAGACTATTCTATCTGATGGCCTTACGAGAATAAATATTGCATTTAATTTTCTAACTGAT
CTCCCTGGATGCAGTCATTTTCACTAGACTTTTCACTTCTGAACATTGAAATGAACCTTCTCT
CAGCCCATCTCTGGATTTTGTTCAGAGCTGCCAGGAAGTTAAACTCTAAATGCGGGAAGAA
ATCCATTCGGGTGTACCTGTGAATTAATAAAATTTTCACTCAGCTTGAAACATATTCAGAGGTC
ATGATGGTTGGATGGTCAGATTCATACACCTGTGAATACCCTTTAAACCTAAGGGGAAGTAG
GTTAAAAGACGTTTCTATCTCCACGAATTATCTTGCAACACAGCTCTGTTGATTGTACCATTTG
TGGTTATTATGCTAGTTCTGGGGTTGGCTGTGGCCTTCTGCTGTCTCCACTTTGATCTGCC
TGGTATCTCAGGATGCTAGGTCAATGCACACAAACATGGCACAGGGTTAGGAAAACAACCCA
AGAACAACCTCAAGAGAAATGTCCGATTCCACGCATTTATTTTATACAGTGAACATGATTCTC
TGTGGGTGAAGAATGAATTGATCCCCAATCTAGAGAAGGAAGATGGTTCTATCTTGATTGTC
CTTTATGAAAGCTACTTTTGACCCTGGCAAAAGCATTAGTGAAAATATTGTAAGCTTCATTGA
GAAAAGCTATAAGTCCATCTTTGTTTTGTCTCCCAACTTTGTCCAGAATGATGGTGCCATT
ATGAATTTCTACTTTGCCCCACCACAATCTCTTCATGAAAATTTCTGATCATATAATTCTTATC
TTACTGGAACCCATTCCATTCTATTGCATTCCCACCAGGTATCATAACTGAAAGCTCTCCT
GGAAAAAAAGCATACTTGAATGGCCCAAGGATAGGCGTAAATGTGGGCTTTTCTGGGCAA
ACCTTCGAGCTGCTATTAATGTTAATGTATTAGCCACCAGAGAAATGTATGAAGTGCAGACA
TTCACAGAGTTAAATGAAGAGTCTCGAGGTTCTACAATCTCTCTGATGAGAACAGATTGTCT
ATAAATCCCACAGTCCTTGGGAAGTTGGGGACCACATACACTGTTGGGATGTACATTGATA
CAACCTTTTATGATGGCAATTTGACAATATTTATTAATAAATAAAAAATGGTTATTCCTTCATA
TCAGTTTCTAGAAGGATTTCTAAGAATGTATCCTATAGAAACACCTTCACAAGTTTATAAGG
GCTTATGGAAAAGGTGTTTCATCCCAGGATTGTTTATAATCATGAAAAATGTGGCCAGGTGC
AGTGGCTCACTCTTGTAAATCCCAGCACTATGGGAGGCCAAGGTGGGTGACCACGAGGTCAA
GAGATGGAGACCATCCTGGCCAACATGGTGAAACCTGTCTCTACTAAAAATACAAAAATTA
GCTGGGCGTGATGGTGCACGCCTGTAGTCCCAGCTACTTGGGAGGCTGAGGCAGGAGAATCG
CTTGAACCCGGGAGGTGGCAGTTGCAGTGAGCTGAGATCGAGCCACTGCACTCCAGCCTGGT
GACAGAGCGAGACTCCATCTCAAAAAAAGAAAAAAGAAAAAAGAAAAAATGGAAAAACATCC
TCATGGCCACAAAATAAGGTCTAATTCATAAATATAGTACATTAATGTAATATAATATTA
CATGCCACTAAAAAGAATAAGGTAGCTGTATATTTCTGGTATGGAAAAACATATTAATAT
GTTATAAACTATTAGGTTGGTGCAAAACTAATTGTGGTTTTTGGCATTGAAATGGCATTGAA
ATAAAAGTGTAAGAAATCTATACCAGATGTAGTAACAGTGGTTTGGGTCTGGGAGGTGGGA
TTACAGGGAGCATTTGATTTCTATGTTGTGATTTCTATAATGTTTGAATTGTTTAGAATGA
ATCTGTATTTCTTTTATAAGTAGAAAAAATAAAGATAGTTTTTACAGCCT

FIGURE 34

MRLIRNIYIFCSIIVMTAEGDAPELPEERELMTNCSNMSLRKVPADLTPATTTLDLSYNLLFQ
LQSSDFHSVSKLRVLILCHNRIQQDLKTFFFNKELRYLDLSNNRLKSVTWYLLAGLRYLDL
SFNDFDTMPICEEAGNMSHLEILGLSGAKIQKSDFQKIAHLHLNNTVFLGFRTLPHYEEGSLP
ILNTTKLHIVLPMDTNFWVLLRDGIKTSKILEMTNIDGKSQFVSYEMQRNLSLENAKTSVLL
LNKVDLLWDDLELILQFVWHTSVEHFQIRNVTFGGKAYLDHNSFDYSNTVMRTIKLEHVHFR
VFYIQQDKIYLLLTKMDIENLTISNAQMPHMLFPNYPTKFQYLNFANNILTDELFKRTIQLP
HLKTLILNGNKLETLSLVSCFANNTPLEHLDLSQNLLQHKNDENC SWPETVVMNLSYNKLS
DSVFRCLPKSIQILD LNNNQIQTVPKETIHLMALRELNIAFNFLTDLPGCSHFSRLSVLNIE
MNFILSPSLDFVQSCQEVKTLNAGRNPFRCCTCELKNFIQLETYSEVMMVGWSDSYTCEYPLN
LRGTRLKDVHLHELSCNTALLIVTIVVIMLVGLAVAFCCCLHFDLPWYLRMLGQCTQTWHRV
RKTTOEQQLKRNVRFHAFISYSEHDSLWVKNELIPNLEKEDGSILICLYESYFDPGKSISENI
VSFIEKSYKSIFVLSPNFVQNEWCHYEFYFAHNNLFHENS DHIILILLEPIPFYCIPTRYHK
LKALLEKKAYLEWPKDRRKCGLEFWANLRAAINVNVLATREMYELQTFTELNEESRGSTISLM
RTDCL

FIGURE 35

GGGGGCTTTCTTGGGCTTGGCTGCTTGGAAACACCTGCCTCCAAGGACCGGCTCGGAGGGGTGCGCGGGAAAG
 GAGGGAAGAAGGAAGGGCGGGGCCGGCCCCCTGCGCCCCGCGCCTCTGCGCGCCCCCTGTCCGCCCCGGC
 CCAGCCCAGCCCAGCCCCGCGGGCCGGTCACACGCGCAGCCAGCCGGCCGCTCCCGCGCCCAAGCGCGCCGCT
 CTGCTGTGCCCTGCGCCCTTGGCCCCGCGCCAGCTTCTGCGCCCCGAGCCCGCCCGGCGCCCCCGGTGACCGTGA
 CCCTGCCCTGGGCGCGGGGCGGAGCAGGCATGCTCCCGCCCCGGGACCGCTACCCCAGCGCTGGCCCTGGTGCTC
 CTGGCAGTGACCTGGCCGGGGTTCGGAGCCAGGGGCGCAGCCCTCGAGGACCGTGATTATTACGGGCGAGGAGAT
 CTGGAGCCGGGAGCCCTACTACGCGCGCCCGGAGCCCGAGCTCGAGACCTTCTCTCCGCCGCTGCCTGCGGGGC
 CCGGGGAGGAGTGGGAGCGGCGCCCGCAGGAGCCCGAGCCGCCCCAAGAGGGCCACCAAGCCCAAGAAAGCTCCC
 AAGAGGGAGAAGTCGGCTCCGGAGCCGCTCCACCAGGTAAACACAGCAACAAAAAGTTATGAGAACCAAGAG
 CTCTGAGAAGGCTGCCAACGATGATCACAGTGTCCGTGTGGCCCGTGAAGATGTGAGAGAGAGTTGCCACCTC
 TTGGTCTGGAAACCTTAAAAATCACAGACTTCCAGTCCATGCCTCCACGGTGAAGCGCTATGGCTGGGGGCA
 CATCGAGGGAGACTCAACATCCAGGCGGGCATTAAATGAAAATGATTTTATGACGGAGCGTGGTGCGCGGGAAG
 AAATGACCTCCAGCAGTGGATTGAAGTGGATGCTCGGCGCCTGACCAGATTCACTGGTGTCACTCAAGGGA
 GGAATCCCTCTGGCTGAGTGACTGGGTGACATCCTATAAGGTGATGGTGAGCAATGACAGCCACACGTGGGT
 ACTGTTAAGAATGGATCTGGAGACATGATATTTGAGGGAAACAGTGAGAAGGAGATCCCTGTTCTCAATGAGCT
 ACCCGTCCCCATGGTGGCCCGCTACATCCGCATAAACCCCTCAGTCTGGTTTGATAATGGGAGCATCTGCATGA
 GAATGGAGATCCTGGGCTGCCCCACTGCCAGATCCTAATAATTATTATCACCGCCGGAACGAGATGACCACCACT
 GATGACCTGGATTTTAAGCACCACAATTATAAGGAAATGCGCCAGTTGATGAAAGTTGTGAATGAAATGTGTCC
 CAATATCACCAGAATTTACAACATTGAAAAAAGCCACCAGGGCCTGAAGCTGTATGCTGTGGAGATCTCAGATC
 ACCCTGGGGAGCATGAAGTCGGTGAGCCCGAGTTCCACTACATCGCGGGGGGCCACGGCAATGAGGTGCTGGGC
 CGGGAGCTGCTGCTGCTGCTGGTGCGAGTTCGTGTGTGAGGAGTACTTGGCCCGGAATGCGCGCATCGTCCACCT
 GGTGGAGGAGACGCGGATTACGTCCTCCCTCCCTCAACCCCGATGGCTACGAGAAGGCCTACGAAGGGGGCT
 CGGAGCTGGGAGGCTGGTCCCTGGGACGCTGGACCCACGATGGAATTGACATCAACAACAACCTTCCCTGATTTA
 AACACGCTGCTCTGGGAGGCAGAGGATCGACAGAATGTCCCCAGGAAAGTTCCCAATCACTATATTGCAATCCC
 TGAGTGGTTTCTGTGCGAAAATGCCACGGTGGCTGCCGAGACCAGAGCAGTCATAGCCTGGATGGAATAACTCC
 TTTTGTGCTGGGCGGCAACCTCGAGGCGGGCAGCTGGTGGTGGCTATCCCTACGACCTGGTGCGGTCCCCC
 TGGAAGACGCAGGAACACACCCCCCAGTACCACGTGTTCCGCTGGCTGGCTACTCCTATGCCTCCAC
 ACACCGCCTCATGACAGACGCCCCGAGGAGGGGTGTGCCACACGGAGGACTTCCAGAAGGAGGAGGGCACTGTCA
 ATGGGGCCTCCTGGCACACCGTGCCTGGAAGTCTGAACGATTTGAGTACCTTCATACAACTGCTTCGAACCTG
 TCCATCTACGTGGGCTGTGATAAATACCCACATGAGAGCCAGCTGCCCGAGGAGTGGGAGAATAACCGGGAATC
 TCTGATCGTGTTTCATGGAGCAGGTTTCATCGTGGCATTAAAGGCTTGGTGAGAGATTACATGGAAAAGGAATCC
 CAAACGCCATTATCTCCGTAGAAGGCATTAACCATGACATCCGAACAGCCAACGATGGGATTACTGGCGCCTC
 CTGAACCCCTGGAGAGTATGTGGTTCACAGCAAAGGCCGAAGGTTTCACTGCATCCACCAAGAACTGTATGGTTGG
 CTATGACATGGGGGCCACAAGGTGTGACTTCACACTTAGCAAAAACCAACATGGCCAGGATCCGAGAGATCATGG
 AGAAGTTTGGGAAGCAGCCCGTCAGCCTGCCAGCCAGGCGGCTGAAGCTGCGGGGGCGGAAGAGACGACAGCGT
 GGGTGAACCTCCTGGGCCCTTGAGACTCGTCTGGGACCCATGCAATTAACCAACCTGGTAGTAGCTCCATAG
 TGGACTCACTCACTGTTGTTTCTCTGTAATTCAAGAAGTGCCTGGAAGAGAGGGTGCATTGTGAGGCAGGTCC
 CAAAAGGGAAGGCTGGAGGCTGAGGCTGTTTTCTTTTCTTTGTTCCCATTTATCCAAATAACTTGGACAGAGCA
 GCAGAGAAAAGCTGATGGGAGTGAGAGAACTCAGCAAGCCAACCTGGGAATCAGAGAGAGAAGGAGAAGGAGGG
 GAGCCTGTCCGTTAGAGCCTCTGGCTGCATAGAAAAGGATTCTGGTGCTTCCCCTGTTTGGCTGGCAGCAAGG
 GTTCCACGTGCATTTGCAATTTGCACAGCTAAAATTGCAGCATTTCGCCAGCTGGGCTGTCCCAAATGTTACCA
 TTTGAGATGCTCCAGGCGTCCTAAGAGAATCCACCCTCTCTGGCCCTGGGACATTGCAAGCTGCTACAAATAA
 ATTCTGTGTTCTTTTGACAATAGCGTCATTGCCAAGTGCACATCAGTGAGCCTCTTGAATCTGTTTGTGCTCCT
 TTTTCAACAAAGGAGTGTGTTTCAAAAAGGAGAGAGAGGCTGAGATCATTGAGGAGTTTGTGGGCGAGCAAGCA
 TGGAGCTTCTTGACAAATTCTGGGTCCATAACAACCCCAAGTCCCTGCTGATCCAGTAGCCCTGGAGGTT
 CCCCAGGTAGGGAGAGCCAGAGGTGCCAGCCTTCTGAAGGGCCAGAAAATTTAGCCTGGATCTCCTCTTTTAC
 CTGCTAGGACTGGAAGAGCCAGAAGTGGGTGGCCTGAAGCCCTCTCTCTGCTTGAGGTATTGCCCTGTGTG
 GAATTGAGTGCTCATGGGTGGCCTCATATCAGCCTGGGAGTTATTTTATATGTAGAATGCCAGATCTTCCA
 GATTAGGCTAAATGTAATGAAAACCTCTTAGGATTATCTGTGGAGCATCAGTTTGGGAAGAATTGAAATTAT
 CTTGCAAGAAAAAGTATGTCTCACTTTTTGTTAATGTGTGCTGCCTCATTGACCTGGGAAAAATGAAAAAAA
 AATAAAGCAATGGTAAGACCTTAAAAAAA

FIGURE 36

MSRPGTATPALALVLLAVTLAGVGAQGALEDPDYYGQEIWSREPYARPEPELETFSPPLP
AGPGEWEERRPQEPRPPKRATKPKKAPKREKSAPEPPPPGKHSNKKVMRTKSSEKAANDDHS
VRVAREDVRESCPPLGLETLKITDFQLHASTVKRYGLGAHRGRLNIQAGINENDFYDGAWCA
GRNDLQQWIEVDARRLTRFTGVITQGRNSLWLSDWVTSYKVMVSNDSTWVTVKNGSGDMIF
EGNSEKEIPVLNELPVPVMVARYIRINPQSWFDNGSICMRMEILGCPLPDPNNYYHRRNEMTT
TDDLDFKHHNYKEMRQLMKVVNEMCPNITRIYNIGKSHQGLKLYAVEISDHPGEHEVGEPEF
HYIAGAHGNEVLGRELLLLLVQFVCQEYLARNARIVHLVEETRIHVLPSLNPDGYEKAYEGG
SELGGWSLGRWTHDGIDINNNFPDLNTLLWEAEDRQNVPRKVPNHYIAIPEWFLSENATVAA
ETRAVIAWMEKIPFVLGGNLQGGELVVAYPYDLVRSPWKTQEHTPTPDDHVFRWLAYSAST
HRLMTDARRRVCHTEDFQKEEGTVNGASWHTVAGSLNDFSYLHTNCFELSIYVGCDKYPHES
QLPEEWENNRESLIVFMEQVHRGIKGLVRDSHGKGIPNAIISVEGINHDIRTANDGDYWRL
NPGEYVVTAKAEGFTASTKNCMVGYDMGATRCDFTLSTKNMARIREIMEKFGKQPVSLPARR
LKLGRKRQRG

FIGURE 37

CTAAGAGGACAAGATGAGGCCCCGGCCTCTCATTTCTCCTAGCCCCCTTCTGTTCTTCCTTGGCCAAAGCTGCAGGGG
 ATTTGGGGGATGTGGGACCTCCAATTTCCCAGCCCCGGCTTCAGCTCTTTCCCAGGTGTTGACTCCAGCTCCAGC
 TTCAGCTCCAGCTCCAGGTGCGGCTCCAGCTCCAGCCGACGCTTAGGCAGCGGAGGTTCTGTGTCCCAGTTGTT
 TTCCAATTTACCCGGCTCCGTGGATGACCGTGGGACCTGCCAGTGTCTGTTCCTGCCAGACACCACCTTTC
 CCGTGGACAGAGTGAACGCTTGAATTCACAGCTCATGTTCTTCTCAGAAAGTTTGAGAAAGAACTTTCTAAA
 GTGAGGGAATATGTCCAATTAATTAGTGTGTATGAAAAAGAACTGTTAAACCTAACTGTCCGAATTGACATCAT
 GGAGAAGGATACCATTTCTTACACTGAACCTGGACTTCGAGCTGATCAAGCTAGAAGTGAAGGAGATGGAAAAAC
 TGGTCATACAGCTGAAGGAGAGTTTTGGTGGAAGCTCAGAAATTGTTGACCAGCTGGAGGTGGAGATAAGAAAT
 ATGACTCTCTTGGTAGAGAAGCTTGAGACACTAGACAAAAACAATGTCCTTGCCATTGCGCCGAGAAATCGTGGC
 TCTGAAGACCAAGCTGAAAGAGTGTGAGGCCCTCTAAAGATCAAAACACCCCTGTGCTCCACCCTCCTCCCACTC
 CAGGGAGCTGTGGTCATGGTGGTGTGGTGAACATCAGCAAACCGTCTGTGGTTTCAGCTCAACTGGAGAGGGTTT
 TCTTATCTATATGGTGCTTGGGGTAGGGATTACTCTCCCCAGCATCCAAACAAAGGACTGTATTGGGTGGCGCC
 ATTGAATACAGATGGGAGACTGTTGGAGTATTATAGACTGTACAACACACTGGATGATTGTGCTATTGTATATAA
 ATGCTCGAGAGTTGCGGATCACCTATGGCCAAGGTAGTGGTACAGCAGTTTACAACAACAACATGTACGTCAAC
 ATGTACAACACCGGGAATATTGCCAGAGTTAACCTGACCACCAACACGATTGCTGTGACTCAAACCTCTCCCTAA
 TGCTGCCTATAATAACCGCTTTTTCATATGCTAATGTTGCTTGGCAAGATATTGACTTTGCTGTGGATGAGAATG
 GATTGTGGGTATTATTTCAACTGAAGCCAGCACTGGTAACATGGTGATTAGTAAACTCAATGACACCACACTT
 CAGGTGCTAAACACTTGGTATACCAAGCAGTATAAACCATCTGCTTCTAACGCCTTCATGGTATGTGGGGTTCT
 GTATGCCACCCGTACTATGAACACCAGAACAGAAGAGATTTTTTACTATTATGACACAACACAGGGAAAGAGG
 GCAAACCTAGACATTGTAATGCATAAGATGCAGGAAAAAGTGCAGAGCATTAACATAACCCCTTTTGACCAGAAA
 CTTTATGTCTATAACGATGGTTACCTTCTGAATTATGATCTTTCTGTCTTGCAGAGCCCCAGTAAAGCTGTTTA
 GGAGTTAGGGTGAAGAGAAAAATGTTTGTGAAAAAATAGTCTTCTCCACTTACTTAGATATCTGCAGGGGTGT
 CTAAAAGTGTGTTTCAATTTTGCAGCAATGTTTAGGTGCATAGTTCTACCACACTAGAGATCTAGGACATTTGTCT
 TGATTTGGTGAGTTCTCTTGGGAATCATCTGCCTCTTCAGGCGCATTTTGCAATAAAGTCTGTCTAGGGTGGGA
 TTGTGAGAGGTCTAGGGGCACTGTGGGCCTAGTGAAGCCTACTGTGAGGAGGCTTCACTAGAAGCCTTAAATTA
 GGAATTAAGGAACTTAAACTCAGTATGGCGTCTAGGGATTCTTTGTACAGGAAATATTGCCCAATGACTAGTC
 CTCATCCATGTAGCACCCTAATTCTTCCATGCCTGGAAGAAACCTGGGCACTTAGTTAGGTAGATTAATATCT
 GGAGCTCCTCGAGGGACCAATCTCCAACCTTTTTTTTCCCTCACTAGCACCTGGAATGATGCTTTGTATGTGG
 CAGATAAGTAAATTTGGCATGCTTATATATTCTACATCTGTAAAGTGCTGAGTTTTATGGAGAGAGGCCCTTTTT
 ATGCATTAAATTTGTACATGGCAAATAAATCCCAGAAGGATCTGTAGATGAGGCACCTGCTTTTTCTTTCTCTC
 ATTTGTCCACCTTACTAAAAGTCAGTAGAATCTTCTACCTCATAACTTCCTTCAAAGGCAGCTCAGAAGATTAG
 AACCAGACTTACTAACCAATTCCACCCCCCACCACCCCTTCTACTGCCTACTTTAAAAAAATTAATAGTTTT
 CTATGGAACTGATCTAAGATTAGAAAAATTAATTTTCTTAAATTTCAATTATGGACTTTTATTTACATGACTCTA
 AGACTATAAGAAAATCTGATGGCAGTGACAAAGTGCTAGCATTTATTGTTATCTAATAAAGACCTTGGAGCATA
 TGTGCAACTTATGAGTGTATCAGTTGTTGCATGTAATTTTTGCCTTTGTTTAAGCCTGGAACCTGTAAGAAAAT
 GAAAATTTAATTTTTTTTTCTAGGACGAGCTATAGAAAAGCTATTGAGAGTATCTAGTTAATCAGTGCAGTAGT
 TGGAAACCTTGCTGGTGTATGTGATGTGCTTCTGTGCTTTTGAATGACTTTATCATCTAGTCTTTGTCTATTTT
 TCCTTTGATGTTCAAGTCTAGTCTATAGGATTGGCAGTTTAAATGCTTTACTCCCCCTTTTAAATAAATGAT
 TAAAATGTGCTTTGAAAAAATAAAAAAAAAAAAAAAAAAAAAA

FIGURE 38

MRPGLSFL LALLFFLGQAAGDLGDVGPPIPSPGFSSFP
GVDSSSSSFSSSSSRSGSSSSRSLGS
GGSVSQLF SNFTGSVDDRGT CQCSVSLPDTTFPVDR
VERLEFTAHVLSQKF EKELSKVREYV
QLISVYEKLLNLTVRIDIMEKDTISYTELDFELIKVE
VKEMEKLVIQLKESFGGSSEIVDQ
LEVEIRNMTLLVEKLETLDKNNVLAIRREIVALKT
KLKECEASKDQNTPVVHPPPTPGSCGH
GGVVNISKPSVVQLNWRGFSYLYGAWGRDYS
PQHHPNKGLYWVAPLNTDGRLL EYYRLYNTLD
DLLLYINARELRITYGQSGTAVYNNNMYVNM
YNTGNIARVNLTNTIAVTQTLPNAAAYNNR
FSYANVAWQDIDFAVDENGLWVIYSTEASTG
NMVISKLNDTTLQVLNTWYTKQYKPSASNAF
MVCGLYATRMTNTRTEEIIFYYYDTNTGKEG
KLDIVMHKMQEKVQSINYNPFDQKLYVYNDG
YLLNYDLSVLQKPQ

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 39

GCTCTGAAGACCAAGCTGAAAGAGTGTGAGGCCTCTAAAGATCAAACACCCCCTGTCGTCCAC
CCTCCTCCCCTCCAGGGAGCTGTGGTCATGGTGGTGTGGTGAACATCAGCAAACCGTCTGT
GGTTCAGCTCAACTGGAGAGGGTTTTCTTATCTATATGGTGCTTGGGGTAGGGATTACTCTC
CCCAGCATCCAAACAAAGGNATGTATTGGGNGGCGCCATTGAATACAGATGGGAGACTGTTG
GAGTATTATAGACTGTACAACCCACTGGATGATTTGCTATTGTATATAAATGCTCGAGAGTT
GCGGATCACCTATGGCCAAGGTAGTGGTACAGCAGTTTACAACAACATGTACGTCAACA
TGTACAACACCGGGNATATTGCCAGAGTTAACCTGACC

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 41

MGSVLGLCSMASWIPCLCGSAPCLLCRCCPSGNNSTVTRLIYALFLLVGVCVACVMLIPGME
EQLNKIPGFCENEKGVVPCNILVGYKAVYRLCFGLAMFYLLLSLLMIKVKSSSDPRAAVHNG
FWFFKFAAAIAIIIGAFFIPEGTFTTVWFYVGMAGAFCFILIQLVLLIDFAHSWNESWVEKM
EEGNSRCWYAALLSATALNYLLSLVAIVLFFVYYTHPASCSENKAFISVNMLLCVGASVMSI
LPKIQESQPRSGLLQSSVITVYTMYLTSAMTNEPETNCNPSLLSIIGYNTTSTVPKEGQSV
QWWHAQGIIGLILFLLCVFYSSIRTSNNSQVNKLTLTSDESTLIEDGGARSDGSLEDGDDVH
RAVDNERDGVTSYSFFHFMLFLASLYIMMTLTNWSRYEPSREMKSQWTAVVVKISSSWIGI
VLYVWTLVAPLVLTNRDFD

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 46

CTCGGGCGCGCACAGGCAGCTCGGTTTGCCCTGCGATTGAGCTGCGGGTTCGCGGCCGCGCCGGCCTCTCCAAT
 GGCAAAATGTGTGTGGCTGGAGGCGAGCGCGAGGCTTTTCGGCAAAGGCAGTCGAGTGTTTGAGACCGGGGCGAG
 TCCTGTGAAAGCAGATAAAAAGAAAACATTTATTAACGTGTCTATTACGAGGGGAGCGCCCGGGGCTGTGCG
 ACTCCCCGCGGAACATTTGGCTCCCTCCAGCTCCGAGAGAGGAGAAGAAGAAAGCGGAAAAGAGGCAGATTAC
 GTCGTTTCCAGCCAAGTGGACCTGATCGATGGCCCTCCTGAATTTATCACGATATTTGATTTATTAGCGATGCC
 CCCTGGTTTGTGTGTACGCACACACACGTGCACACAAGGCTCTGGCTCGCTTCCCTCCCTCGTTTCCAGCTCC
 TGGGCGAATCCACATCTGTTTCAACTCTCCGCCGAGGGGCGAGCAGGAGCGAGAGTGTGTGCAATCTCGAGTG
 AAGAGGGACGAGGGGAAAAGAAAACAAAGCCACAGACGCAACTTGAGACTCCCCGCATCCCCAAAAGAACAGAT
 CAGCAAAAAAGAAAGATGGGGCCCCCGAGCCTCGTGCTGTGCTTGTGTCCGCAACTGTGTTCTCCCTGCTGGG
 TGGAAGCTCGGCCCTCCTGTGCGCACCACCGCTGAAAGGCAGGTTTCAGAGGGACCGCAGGAACATCGCCCCA
 ACATCATCTCTGGTGTGACGGACGACCAGGATGTGGAGCTGGGTTCCATGCAGGTGATGAACAAGACCCGGCGC
 ATCATGGAGCAGGGCGGGGCGCACCTTCATCAACGCCCTTCGTGACCACACCCATGTGCTGCCCCCTCACGCTCCTC
 CATCCTCACTGGCAAGTACGTCCACAACCACAACACCTACACCAACAATGAGAACTGCTCCTCGCCCTCCTGGC
 AGGCACAGCAGAGAGCCGCACCTTTGCCGTGTACCTCAATAGCACTGGCTACCGGACAGCTTCTTCGGGAAG
 TATCTTAATGAATACAACGGCTCCTACGTGCCACCCGGCTGGAAGGAGTGGGTTCGGACTCCTTAAAAACTCCCG
 CTTTTATAACTACACGCTGTGTGCGAACGGGGTGAAGAGAAGCACGGCTCCGACTACTCCAAGGATTACCTCA
 CAGACCTCATACCAATGACAGCGTGAGCTTCTCCGCACGTCCAAGAAGATGTACCCGCACAGGCCAGTCTC
 ATGGTCATCAGCCATGCAGCCCCCACGGCCCTGAGGATTACGCCCCACAATATTCACGCCCTCTTCCCAAACGC
 ATCTCAGCACATCAGCCGAGCTACAACCTACGCGCCCAACCCGGACAAACACTGGATCATGCGCTACACGGGGC
 CCATGAAGCCCATCCACATGGAATTCACCAACATGCTCCAGCGGAAGCGCTTGCAGACCCCTCATGTGCGTGGAC
 GACTCCATGGAGACGATTTACAACATGCTGGTTGAGACGGGGCAGCTGGACAACACGTACATCGTATACACCGC
 CGACCACGGTTACCACATCGGCCAGTTTGGCCTGGTGAAGGGAAATCCATGCCATATGAGTTTGACATCAGGG
 TCCCGTTCTACGTGAGGGGGCCCCAACGTGGAAGCCGGCTGTCTGAATCCCCACATCGTCTCAACATTGACCTG
 GCCCCACCATCCTGGACATTGCAGGCCTGGACATACCTGCGGATATGGACGGGAAATCCATCCTCAAGCTGCT
 GGACACGGAGCGGCCGGTGAATCGGTTTCACTTGAAAAAGAAATGAGGGTCTGGCGGGACTCCTTCTTGGTGG
 AGAGAGGCAAGCTGTACACAAGAGAGACAATGACAAGGTGGACGCCAGGAGGAGAACTTCTGCCCAAGTAC
 CACGTGTGAAGGACCTGTGTACGCGTGTGAGTACCAGACGGCGTGTGAGCAGCTGGGACAGAAGTGGCAGTG
 TGTGGAGGACGCCACGGGGAAGCTGAAGCTGCATAAGTGAAGGGCCCCATGCGGCTGGGCGGCAGCAGAGCCC
 TCTCCAACCTCGTGCCCAAGTACTACGGGCAGGGCAGCGAGGCCTGCACCTGTGACAGCGGGGACTACAAGCTC
 AGCCTGGCCGGACGCCGAAAAAACTCTTCAAGAAGAAGTACAAGGCCAGCTATGTCCGCAGTCGCTCCATCCG
 CTCAGTGGCCATCGAGGTGGACGGCAGGGTGTACCACGTAGGCCTGGGTGATGCCGCCAGCCCCGAAACCTCA
 CCAAGCGGCACTGGCCAGGGGCCCCCTGAGGACCAAGATGACAAGGATGGTGGGACTTCTTCCAACTGGAGGC
 CTTCCCGACTACTCAGCCGCCAACCCCATTAAGTGAACATCGGTGCTACATCCTAGAGAACGACACAGTCCA
 GTGTGACCTGGACCTGTACAAGTCCCTGCAGGCCTGGAAGACCACAAGCTGCACATCGACCACGAGATTGAAA
 CCCTGCAGAACAAAATTAAGAACCTGAGGGAAGTCCGAGGTACCTGAAGAAAAAGCGGCCAGAAGAAATGTGAC
 TGTACAAAATCAGCTACCACACCCAGCACAAAGGCCGCCCTCAAGCACAGAGGCTCCAGTCTGCATCCTTTTCA
 GAAGGGCCTGCAAGAGAAGGACAAGGTGTGGCTGTTGCGGGAGCAGAAGCGCAAGAAGAAACTCCGCAAGCTGC
 TCAAGCGCCTGCAGAACACGACAGCTGCAGCATGCCAGGCCTCACGTGCTTCACCCAGACAACCAGCACTGG
 CAGACGGCGCCTTTCTGGACACTGGGGCCTTTCTGTGCTGCACCAGCGCCAACAATAACACGTACTGGTGCAT
 GAGGACCATCAATGAGACTCACAATTTCTCTTCTGTGAATTTGCAACTGGCTTCTTAGAGTACTTTGATCTCA
 ACACAGACCCCTACCAGCTGATGAATGCAGTGAACACACTGGACAGGGATGTCTCAACCAGCTACACGTACAG
 CTCATGGAGCTGAGGAGCTGCAAGGGTTACAAGCAGTGTAAACCCCGGACTCGAAACATGGACCTGGATGGAGG
 AAGCTATGAGCAATACAGGCAGTTTCAGCGCTCGAAAGTGGCCAGAAATGAAGAGACCTTCTTCCAAATCACTGG
 GACAACCTGTGGGAAGGCTGGGAAGGTTAAGAAACAACAGAGGTGGACCTCCAAAACATAGAGGCATCACCTGA
 CTGCACAGGCAATGAAAAACCATGTGGGTGATTTCCAGCAGACCTGTGCTATTGGCCAGGAGGCCTGAGAAAGC
 AAGCACGCACTCTCAGTCAACATGACAGATTCTGGAGGATAACCAGCAGGAGCAGAGATAACTTCAGGAAGTCC
 ATTTTTGCCCTGCTTTTGTCTTTGGATTATACCTCACCAGCTGCACAAAATGCATTTTTTTCGTATCAAAAAGTC
 ACCACTAACCTCCCCAGAAAGCTCACAAAGGAAAACGGAGAGAGCGAGCGAGAGAGATTTCCTTGGAAATTC
 TCCCAAGGGCGAAAGTCATTGGAATTTTTAAATCATAGGGGAAAAGCAGTCTGTCTAAATCCTCTTATCTT
 TTGGTTTGTACAAAGAAGGAACATAAGAAGCAGGACAGAGGCAACGTGGAGAGGCTGAAAACAGTGCAGAGACG
 TTTGACAAATGAGTCAGTAGCACAAAAGAGATGACATTTACCTAGCACTATAAACCTGGTTGCCTCTGAAGAAA
 CTGCGCTTCATTGTATATATGTGACTATTTACATGTAATCAACATGGGAACCTTTTAGGGGAACCTAATAAGAAAT
 CCCAATTTTCAGGAGTGGTGGTGTCAATAACGCTCTGTGGCCAGTGTAAAAGAAAAA

FIGURE 48

AACAAAGTTCAGTGACTGAGAGGGCTGAGCGGAGGCTGCTGAAGGGGAGAAAGGAGTGAGGA
GCTGCTGGGCAGAGAGGGGACTGTCCGGCTCCCAGATGCTGGGCCTCCTGGGGAGCACAGCCC
TCGTGGGATGGATCACAGGTGCTGCTGTGGCGGTCTGCTGCTGCTGCTGCTGCTGGCCACC
TGCCTTTTCCACGGACGGCAGGACTGTGACGTGGAGAGGAACCGTACAGCTGCAGGGGAAA
CCGAGTCCGCCGGGCCCAGCCTTGGCCCTTCCGGCGGGGGGCCACCTGGGAATCTTTCACC
ATCACCGTCATCCTGGCCACGTATCTCATGTGCCGAATGTGGGCCTCCACCACCACCACCAC
CCCCGCCACACCCCTCACCACCTCCACCACCACCACCACCCCCACCGCCACCATCCCCGCCA
CGCTCGCTGAGGCTGCTGTGCCGGTGCCTGTGGACAGCAGCTGCCCCCTGCCCTCCCATCTG
TCCCAGGACAAGTGGACCCCATGTTTCCATGTGGAAGGATGCATCTCTGGGGTGAACGAGG
GGAACAATAGACTGGGGCTTGCTCCAGCTGCATTTGCATGGCATGCCCCAGTGTAATATGGC
AGCAGAGAATGGAGGAACACTGGGTCTGCAGTGCTGAAGGGTTTGGGGAGTGGAGAGCAAGG
GTGCTCTTTCGGGGCTGGACAGCCCGTCTTGTGACAGTGAATCCCAGTGAGCCCCAGAAATG
ACAAGCGTGTCTTGGCAGAGCCAGCACACAAGTGGATGTGAAGTGCCCGTCTTGACCTCCTC
ATCAGGCTGCTGCAGGCCTCTGGCGGGCAGGGCACTGGGAGAGGCCCTGAGAATGTCCTTTT
GGTTTGGAGAAGGCAGTGTGAGGCTGCACAGTCAATTATCGGTGCCTTAGTCCAAGAAAAT
AAAAACCACTAAGAAGCTTTAAAAAAAAAAAAAAAAAAAAA

49/330

FIGURE 49

MLGLLGSTALVGWITGAAVAVLLLLLLLLLATCLFHGRQDCDVERNRTAAGGNRVRRRAQWPWFR
RRGHLGIFHHHRHPGHVSHVPNVGLHHHHHPRHTPHHLHHHHHPRHHHPRHAR

49/330

FIGURE 50

GGCGGCTGCTGAGCTGCCTTGAGGTGCAGTGTGGGGATCCAGAGCC**ATG**TCGGACCTGCTA
CTACTGGGCCTGATTGGGGGCTGACTCTCTTACTGCTGCTGACGCTGCTGGCCTTTGCCGG
GTACTCAGGGCTACTGGCTGGGGTGGAAGTGAGTGCTGGGTACCCCCCATCCGCAACGTCA
CTGTGGCCTACAAGTTCCACATGGGGCTCTATGGTGAGACTGGGCGGCTTTTCACTGAGAGC
TGCAGCATCTCTCCCAAGCTCCGCTCCATCGCTGTCTACTATGACAACCCCCACATGGTGCC
CCCTGATAAGTGCCGATGTGCCGTGGGCAGCATCCTGAGTGAAGGTGAGGAATCGCCCTCCC
CTGAGCTCATCGACCTCTACCAGAAATTTGGCTTCAAGGTGTTCTCCTTCCCGGCACCCAGC
CATGTGGTGACAGCCACCTTCCCCTACACCACCATTCTGTCCATCTGGCTGGCTACCCGCCG
TGTCCATCCTGCCTTGACACCTACATCAAGGAGCGGAAGCTGTGTGCCTATCCTCGGCTGG
AGATCTACCAGGAAGACCAGATCCATTTTATGTGCCCCTGGCACGGCAGGGAGACTTCTAT
GTGCCTGAGATGAAGGAGACAGAGTGGAATGGCGGGGGCTTGTGGAGGCCATTGACACCCA
GGTGGATGGCACAGGAGCTGACACAATGAGTGACACGAGTTCTGTAAGCTTGAAGTGAGCC
CTGGCAGCCGGGAGACTTCAGCTGCCACACTGTCACCTGGGGCGAGCAGCCGTGGCTGGGAT
GACGGTGACACCCGCAGCGAGCACAGCTACAGCGAGTCAGGTGCCAGCGGCTCCTCTTTTGA
GGAGCTGGACTTGGAGGGCGAGGGGCCCTTAGGGGAGTCACGGCTGGACCCTGGGACTGAGC
CCCTGGGGACTACCAAGTGGCTCTGGGAGCCCACTGCCCCTGAGAAGGGCAAGGAG**TAA**CCC
ATGGCCTGCACCCTCCTGCAGTGCAGTTGCTGAGGAACTGAGCAGACTCTCCAGCAGACTCT
CCAGCCCTCTTCTCCTTCTCTGGGGGAGGAGGGGTTCTGAGGGACCTGACTTCCCCTGC
TCCAGGCCCTCTTGCTAAGCCTTCTCCTCACTGCCCTTTAGGCTCCCAGGGCCAGAGGAGCCA
GGGACTATTTTCTGCACCAGCCCCCAGGGCTGCCGCCCTGTTGTGTCTTTTTTTCAGACTC
ACAGTGGAGCTTCCAGGACCCAGAATAAAGCCAATGATTTACTTGTTCACCTGGAAAAAAA
AAAAA

FIGURE 51

MSDLLLLGLIGGLTLLLLLTLLAFAGYSGLLAGVEVSAGSPPIRNVTVAYKFHMGLYGETGR
LFTESCSISPKLRSIAVYYDNPHMVPPDKCRCAVGSILSEGEESPSPELIDLYQKFGFKVFS
FPAPSHVVTATFPYTTILSIWLATTRVHPALDTYIKERKLCAYPRLEIYQEDQIHFMCPALAR
QGDFYVPEMKETEWKWRGLVEAIDTQVDGTGADTMSDTSSVSLEVSPGSRETSAA TLSPGAS
SRGWDDGDTRSEHSYSESGASGSSFEELDLEGEGLGESRLDPGTEPLGTTKWLWEPTAPEK
GKE

MSDLLLLGLIGGLTLLLLLTLLAFAGYSGLLAGVEVSAGSPPIRNVTVAYKFHMGLYGETGR
LFTESCSISPKLRSIAVYYDNPHMVPPDKCRCAVGSILSEGEESPSPELIDLYQKFGFKVFS
FPAPSHVVTATFPYTTILSIWLATTRVHPALDTYIKERKLCAYPRLEIYQEDQIHFMCPALAR
QGDFYVPEMKETEWKWRGLVEAIDTQVDGTGADTMSDTSSVSLEVSPGSRETSAA TLSPGAS
SRGWDDGDTRSEHSYSESGASGSSFEELDLEGEGLGESRLDPGTEPLGTTKWLWEPTAPEK
GKE

FIGURE 52

CCGCGGGAACGCTGTCCTGGCTGCCGCCACCCGAACAGCCTGTCTTGGTGCCCCGGCTCCCT
GCCCCGCGCCAGTC**ATG**ACCCTGCGCCCCCTCACTCCTCCCGCTCCATCTGCTGCTGCTGCT
GCTGCTCAGTGCGGCGGTGTGCCGGGCTGAGGCTGGGCTCGAAACCGAAAGTCCCGTCCGGA
CCCTCCAAGTGGAGACCCTGGTGGAGCCCCCAGAACCATGTGCCGAGCCCGCTGCTTTTGGA
GACACGCTTCACATACACTACACGGGAAGCTTGGTAGATGGACGTATTATTGACACCTCCCT
GACCAGAGACCCTCTGGTTATAGAACTTGGCCAAAAGCAGGTGATTCCAGGTCTGGAGCAGA
GTCTTCTCGACATGTGTGTGGGAGAGAAGCGAAGGGCAATCATTCTTCTCACTTGGCCTAT
GGAAAACGGGGATTTCCACCATCTGTCCCAGCGGATGCAGTGGTGCAGTATGACGTGGAGCT
GATTGCACTAATCCGAGCCAACACTACTGGCTAAAGCTGGTGAAGGGCATTCTGCCTCTGGTAG
GGATGGCCATGGTGCCAGCCCTCCTGGGCCTCATTGGGTATCACCTATACAGAAAGGCCAAT
AGACCCAAAGTCTCCAAAAGAAGCTCAAGGAAGAGAAACGAAACAAGAGCAAAAAGAA**TA**
ATAAATAATAAATTTTAAAAAACTTAAAAAAAAAAAAAAAAAAAA

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 53

MTLRPSLLPLHLLLLLLLLLSAAVCRAEAGLETESPVRTLQVETLVEPPEPCAEPAAFCDTLHI
HYTGSLVDGRIIDTSLTRDPLVIELGQKQVIPGLEQSLLDMCVGEKRRAIIPSHLAYGKRGF
PPSVPADAVVQYDVELIALIRANYWLKLVKGILPLVGMAMVPALLGLIGYHLYRKANRPKVS
KKKLKEEKRNKSKKK

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 54

CCCGGGAACGTGTTCCCTGGCTGCCGCACCCGAACAGCCTGTCCTGGTGCCCCGGCTCCCTGC
CCCGCGCCCAAGTCATGACCCTGCGCCCCCTCACTCCTCCCGCTCCATCTGCTGCTGCTGCTGC
TGCTCAGTGCGGCGGTGTGCCGGGCTGAGGCTGGGCTCGAAACCGAAAGTCCCGTCCGGACC
CTCCAAGTGGAGACCCTGGTGGAGCCCCCAGAACCATGTGCCGAGCCCGCTGCTTTTGGAGA
CACGCTTCACATACACTACACGGGAAGCTTGGTAGATGGACGTATTATTGACACCTCCCTGA
CCAGAGACCCTCTGGTTATAGAACTTGGCCAAAAGCAGGTGATTCCAGGTCTGGAGCAGAGT
CTTCTCGACATGTGTGTGGGAGAGAAGCGAAGGGCAATCATTCTTCTCACTTGGCCTATGG
AAAACGGGGATTTCCACCATCTGTCCCAGCGGATGCAGTGGTGCAGTATGACGTGGAGCTGA
TTGCACTAATCCGAGCCAACTACTGGCTAAAGCTGGTGAAGGGCATTTTGCCTCTGGTAGGG
ATGGCCATGGTGCCACCCTCCTGGGCCTCATTGGGTATCACCTATACAGAAAGGCCAATAGA
CCCAAAGTCTCCAAAAAGAAGCTCAAGGAAGAGAAACGAAACAAGAGCAAAAAGAAATAATA
AATAATAAATTTTAAAAAACTTA

FIGURE 55

CCGAAAGTCCCGTCCGGACCCCTCCAAGTGGAGACCCTGGTGGAGCCCCCAGAACCATGTGCC
GAGCCCGCTGCTTTTGGAGACACGCTTCACATACACTACACGGGAAGCTTGGTAGATGGACG
TATTATTGACACCTCCCTGACCAGAGACCCTCTGGTTATAGA ACTTGGCCAAAAGCAGGTGA
TTCCAGGTCTGGAGCAGAGTCTTCTCGACATGTGTGTGGGAGAGAAGCGAAGGGCAATCATT
CCTTCTCACTTGGCCTATGGAAAACGGGGATTTCCACCATCTGTCCCAGCGGATGCAGTGGT
GCAGTATGACGTGGAGCTGATTGCACTAATCCGAGCCAACTACTGGCTAAAGCTGGTGAAGG
GCATTTTGCCTCTGGTAGGGATGGCCATGGTGCCAGCCCTCCTGGGCCTCATTGGGTATCAC
CTATACAGAAAGGCCAATAGACCCAAAGTCTCCAAAAGAAGCTCAAGGAAGAGAAACGAAA
CAAGAGCAAAAAGAAATAATAATAATAAATTTAAAAAACTTAAAA

CCGAAAGTCCCGTCCGGACCCCTCCAAGTGGAGACCCTGGTGGAGCCCCCAGAACCATGTGCC
GAGCCCGCTGCTTTTGGAGACACGCTTCACATACACTACACGGGAAGCTTGGTAGATGGACG
TATTATTGACACCTCCCTGACCAGAGACCCTCTGGTTATAGA ACTTGGCCAAAAGCAGGTGA
TTCCAGGTCTGGAGCAGAGTCTTCTCGACATGTGTGTGGGAGAGAAGCGAAGGGCAATCATT
CCTTCTCACTTGGCCTATGGAAAACGGGGATTTCCACCATCTGTCCCAGCGGATGCAGTGGT
GCAGTATGACGTGGAGCTGATTGCACTAATCCGAGCCAACTACTGGCTAAAGCTGGTGAAGG
GCATTTTGCCTCTGGTAGGGATGGCCATGGTGCCAGCCCTCCTGGGCCTCATTGGGTATCAC
CTATACAGAAAGGCCAATAGACCCAAAGTCTCCAAAAGAAGCTCAAGGAAGAGAAACGAAA
CAAGAGCAAAAAGAAATAATAATAATAAATTTAAAAAACTTAAAA

FIGURE 56

CTGCTGCATCCGGGTGTCTGGAGGCTGTGGCCGTTTTGTTTTCTTGGCTAAAATCGGGGGAG
TGAGGCGGGCCGGCGCGGCGGACACCGGGCTCCGGAACCACTGCACGACGGGGCTGGACTG
ACCTGAAAAAAATGTCTGGATTTCTAGAGGGCTTGAGATGCTCAGAATGCATTGACTGGGGG
GAAAAGCGCAATACTATTGCTTCCATTGCTGCTGGTGTACTATTTTTTACAGGCTGGTGGAT
TATCATAGATGCAGCTGTTATTTATCCCACCATGAAAGATTTCAACCACTCATACCATGCCT
GTGGTGTATAGCAACCATAGCCTTCCTAATGATTAATGCAGTATCGAATGGACAAGTCCGA
GGTGATAGTTACAGTGAAGGTTGTCTGGGTCAAACAGGTGCTCGCATTGCGCTTTTCGTTGG
TTTCATGTTGGCCTTTGGATCTCTGATTGCATCTATGTGGATTCTTTTTGGAGGTTATGTTG
CTAAAGAAAAAGACATAGTATACCCTGGAATTGCTGTATTTTTCCAGAATGCCTTCATCTTT
TTTGGAGGGCTGGTTTTTAAGTTTGGCCGCACTGAAGACTTATGGCAGTGAACACATCTGAT
TTCCACAGCACAACAGCCCTGCATGGGTGTTGTTGTTTTTTACTGCTCACTCCCAACCTT
TTGTAATGCCATTTTCTAACTTATTTCTGAGTGTAGTCTCAGCTTAAAGTTGTGTAATACT
AAAATCACGAGAACACCTAAACAACAACCAAAAATCTATTGTGGTATGCACTTGATTAACCT
ATAAAATGTTAGAGGAACTTTCACATGAATAATTTTTGTCAAATTTTATCATGGTATAATT
TGTA AAAAATAAAAAGAAATTACAAAAGAAATTATGGATTTGTCAATGTAAGTATTTGTCATA
TCTGAGGTCCAAAACCACAATGAAAGTGCTCTGAAGATTTAATGTGTTTATTCAAATGTGGT
CTCTTCTGTGTCAAATGTTAAATGAAATATAAACATTTTTTTAGTTTTTTAAATATTCCGTGG
TCAAATTTCTTCCTCACTATAATTGGTATTTACTTTTACCAAAAATTCTGTGAACATGTAAT
GTA ACTGGCTTTTGAGGGTCTCCCAAGGGGTGAGTGGACGTGTTGGAAGAGAGAAGCACCAT
GGTCCAGCCACCAGGCTCCCTGTGTCCCTTCCATGGGAAGGTCTTCCGCTGTGCCTCTCATT
CCAAGGGCAGGAAGATGTGACTCAGCCATGACACGTGGTTCTGGTGGGATGCACAGTCACTC
CACATCCACCACTG

57/330

FIGURE 57

MSGFLEGLRCSECIDWGEKRNTIASIAAGVLFFTGWIIIDA AVIYPTMKDFNHSYHACGVI
ATIAFLMINAVSNGQVRGDSYSEGCLGQTGARIWLFVGFMLAFGSLIASMWILFGGYVAKEK
DIVYPGIAVFFQNAFIFFGGLVFKFGRTEDLWQ

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 58

TTCTTGGCTAAAATCGGGGGAGTGAGGCGGGCCGGCGCGGCGGACACCGGGCTCCGGAACC
ACTGCACGACGGGGCTGGACTGACCTGAAAAAAATGTCTGGATTTCTAGAGGGCTTGAGATG
CTCAGAAATGCATTGACTGGGGGGAAAAGCGCAATACTATTGCTTCCATTGCTGCTGGTGTAC
TATTTTTTTACAGGCTGGTGGATTATCATAGATGCAGCTGTTATTTATCCCACCATGAAAGAT
TTCAACCACTCATACCATGCCTGTGGTGTATAGCAACCATAGCCTTCCTAATGATTAATGC
AGTATCGAATGGACAAGTCCGAGGTGATAGTTACAGTGAAGGTTGTCTGGGTCAAACAGGTG
CTCGCATTTGGCTTTTCGTTGGTTTCATGTTGGCCTTTGGATCTCTGATTGCATCTATGTGG
ATTCTTTTTGGAGGTTATGTTGCTAAAGAAAAAGACATAGTATACCCTGGAATTGCTGTATT
TTTCCAGAATGCCTTCATCTTTTTTTGGAGGGCTGGTTTTTTAAGTTTGGC

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 60

GGACACCGGGTTCCGGACCAATGCANGACGGGGTGGANTGACCTGAAAAAATGTTTGGATT
TTTAGAGGGCTTGAGATGNTCAGAATGCATTGACTGGGGGAAAAGCGCAATANTATTGCTTT
CCATTGCTGCTGGTGTACTATTTTTTACAGGGTGGTGGATTATCATAGATGCAGCTGTTATT
TATCCCACCATGAAAGATTTNAACCACTCATACCATGCCTGTGGTGTTATAGCAACCATAGC
CTTCCTAATGATTAATGCAGTATCGAATGGACAAGTCCGAGGTGATAGTTACAGTGAAGGTT
GTTTGGGTCAAACAGGTGNTCGCATTGGCTTTTCGTTGGTTTCATGTTGGCCTTTGGATTT
CTGATTGNATTCTATGCGGATTCTTCTTGGAGGTTATGTTGCTAAAGAAAAAGACATAGTAT
ACCCTGGAATTNCTNTATTTTTTCCAGAATGCC

60/330

FIGURE 61

TAGAGGGCTTGAGATGCTCAGAATGCATTGACTGGGGGGAAAAGCGCAATANTATTGCTTCC
ATTGNTGNTGGTGTANTATTTTTTTTACAGGCTGGTGGATTATNATAGATGCAGCTGTTATTT
ATCCCACCATGAAAGATTTNAACCANTCATACCATGCCTGTGGTGTTATAGCAACCATAGCC
TTCCTAATGATTAATGCAGTATNGAATGGACAAGTCCGAGGTGATAGTTACAGTGAAGGTTG
TTTGGGTCAAACAGGTGNTNGCATTGGCTTTTNGTTGGTTTCATGTTGGCCTTTGGATCTN
TGATTGCATTTATGTGGATTNTTTTTTGGAGGTTATGTTGCTAAAGNAAAAGACATAGTATAC
CCTGT

61/330

FIGURE 62

GGGAGGCTGTGNCCGTTTTGTTTTNTTGGCTAAAATCGGGGGAGTGAGGCGGCCCGGCGCGG
CGNGACACCGGGTTCCGGGAACCATTGACGACGGGGTGGACTGACCTGAAAAAATGTTTG
GATTTNTAGAGGGCTTGAGATGCTCAGAATGCATTGACTGGGGGGAAAAGCGCAATACTATT
GCTTCCATTGCTGCTGGTGTACTATTTTTTACAGGCTGGTGGATTATCATAGATGCAGCTGT
TATTTATCCCACCATGAAAGATTTCAACCACTCATACCATGCCTGTGGTGTATAGCAACCA
TAGCCTTCCTAATGATTAATGCAGTATCGAATGGACAAGTCCGAGGTGATAGTTACAGTGAA
GGTTGTCTGGGTCAAACAGGTGCTCGCATTGGCTTTTCGTTGGTTTCATGTTGGCCTTTGG
ATNTCTGATTGCATCTATGTGGATTCTTTTTGGAGGTTATGTTGCTAAAGAAAAAGACATAG
TATACCCTGGAATTGCTGTATTTTTCCAGAATGCCTTCATNTTTTTTGGAGGGCTG

GGGAGGCTGTGNCCGTTTTGTTTTNTTGGCTAAAATCGGGGGAGTGAGGCGGCCCGGCGCGG
CGNGACACCGGGTTCCGGGAACCATTGACGACGGGGTGGACTGACCTGAAAAAATGTTTG
GATTTNTAGAGGGCTTGAGATGCTCAGAATGCATTGACTGGGGGGAAAAGCGCAATACTATT
GCTTCCATTGCTGCTGGTGTACTATTTTTTACAGGCTGGTGGATTATCATAGATGCAGCTGT
TATTTATCCCACCATGAAAGATTTCAACCACTCATACCATGCCTGTGGTGTATAGCAACCA
TAGCCTTCCTAATGATTAATGCAGTATCGAATGGACAAGTCCGAGGTGATAGTTACAGTGAA
GGTTGTCTGGGTCAAACAGGTGCTCGCATTGGCTTTTCGTTGGTTTCATGTTGGCCTTTGG
ATNTCTGATTGCATCTATGTGGATTCTTTTTGGAGGTTATGTTGCTAAAGAAAAAGACATAG
TATACCCTGGAATTGCTGTATTTTTCCAGAATGCCTTCATNTTTTTTGGAGGGCTG

FIGURE 63

FIGURE 64

MWLPLVLLLAVALLLAVLCKVYLGLFSGSSPNPFSEDEVKRPPAPLVTDKEARKKVLKQAFSAN
QVPEKLDVVVIGSGFGGLAAAAILAKAGKRVLVLEQHTKAGGCCHTFGKNGLEFDTGIIHYIG
RMEEGSIGRFILDQITEGQLDWAPLSSPFDIMVLEGPNGRKEYPMYSGEKAYIQGLKEKFPQ
EEAIIIDKYIKLVKVVSSGAPHAILLKFLPLFVVQLLDRCGLLTRFSPFLQASTQSLAEVLQQ
LGASSELQAVLSYIFPTYGVTPNHSAFSMHALLVNHYMKGGFYPRGGSSEIAFHTIPVIQRA
GGAVLTKATVQSVLLDSAGKACGVSVKKGHELVNIYCPIVVSNAGLFNTYEHLLPGNARCLP
GVKQQLGTVRPGMGMTSVFICLRGTKEDLHLPSTNYYVYYDTMDQAMERYVSMFREEAAEH
IPLLFFAFPSAKDPTWEDRFPGRSTMIMLIPTAYEWFEWQAELKGKRGSDYETFKN SFVEA
SMSVVLKLFPPLEGKVESVTAGSPLTNQFYLAAPRGACYGADHDLGRLHPCVMASLRAQSPI
PNLYLTGQDIFTCGLVGALQGALLCSSAILKRNLYS DLKNLDSRIRAQKKKN

[illegible]

FIGURE 66

MRVRIGLTLLLCVLLSLASASSDEEGSQDESLDSKTTLTSDSVKDHTTAGRVVAGQIFLD
SESELESSIQEEEDSLKSQEGESVTEDISFLESPNPENKDYEEPKKVRKPALTAIEGTAHG
EPCHFPLFLDKEYDECTSDGREDGRLWCATTYDYKADEKWGFCETEEEAARRQMGEAEMM
YQTGMKILNGSNKKSQKREAYRYLQKAASMNHTKALERSYALLFGDYLPQNIQAAREMF EK
LTEEGSPKGQTALGFLYASGLGVNSSQAKALVYYTFGALGGNLI AHMVLVSRL

66/330

FIGURE 67

CTTCCCAGCCCTGTGCCCCAAAGCACCTGGAGCATATAGCCTTGCAGAACTTCTACTTGCCT
GCCTCCCTGCCTCTGGCC**ATG**GCCTGCCGGTGCCTCAGCTTCCTTCTGATGGGGACCTTCCT
GTCAGTTTCCCAGACAGTCCTGGCCCAGCTGGATGCACTGCTGGTCTTCCCAGGCCAAGTGG
CTCAACTCTCCTGCACGCTCAGCCCCCAGCACGTCACCATCAGGGACTACGGTGTGTCCTGG
TACCAGCAGCGGGCAGGCAGTGCCCCCTCGATATCTCCTCTACTACCGCTCGGAGGAGGATCA
CCACCGGCCTGCTGACATCCCCGATCGATTCTCGGCAGCCAAGGATGAGGCCCACAATGCCT
GTGTCCTCACCATTAGTCCCGTGCAGCCTGAAGACGACGCGGATTACTACTGCTCTGTTGGC
TACGGCTTTAGTCCC**TAG**GGGGTGGGGTGTGAGATGGGTGCCTCCCCCTCTGCCTCCCATTTCT
GCCCCTGACCTTGGGTCCCTTTTAAACTTTCTCTGAGCCTTGCTTCCCCTCTGTAAAATGGG
TTAATAATATTCAACATGTCAACAAC

CTTCCCAGCCCTGTGCCCCAAAGCACCTGGAGCATATAGCCTTGCAGAACTTCTACTTGCCT
GCCTCCCTGCCTCTGGCCATGGCCTGCCGGTGCCTCAGCTTCCTTCTGATGGGGACCTTCCT
GTCAGTTTCCCAGACAGTCCTGGCCCAGCTGGATGCACTGCTGGTCTTCCCAGGCCAAGTGG
CTCAACTCTCCTGCACGCTCAGCCCCCAGCACGTCACCATCAGGGACTACGGTGTGTCCTGG
TACCAGCAGCGGGCAGGCAGTGCCCCCTCGATATCTCCTCTACTACCGCTCGGAGGAGGATCA
CCACCGGCCTGCTGACATCCCCGATCGATTCTCGGCAGCCAAGGATGAGGCCCACAATGCCT
GTGTCCTCACCATTAGTCCCGTGCAGCCTGAAGACGACGCGGATTACTACTGCTCTGTTGGC
TACGGCTTTAGTCCC**TAG**GGGGTGGGGTGTGAGATGGGTGCCTCCCCCTCTGCCTCCCATTTCT
GCCCCTGACCTTGGGTCCCTTTTAAACTTTCTCTGAGCCTTGCTTCCCCTCTGTAAAATGGG
TTAATAATATTCAACATGTCAACAAC

68/330

FIGURE 68

MACRCLSFLLMGTFLSVSQTVLAQLDALLVFPGQVAQLSCTLSPQHVTIRDYGVSWYQQRAG
SAPRYLLYYRSEEDHHRPADIPDRFSAKDEAHNACVLTISPVPEDDADYYCSVGYGFS

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 69

GCCGCCCCGCCCCGAGACCGGGCCCCGGGGGCGCGGGGCGGCGGGATGCGGCGCCCCGGGGCGG
 CGATGACCGCGGAGCGCACGCCCGCGGGGCCCGCCCTGACCCCGCCGCCCCGCTGAGCCC
 CCGCCGAGGTCCGGACAGGCCGAGATGACGCCGAGCCCCCTGTTGCTGCTCCTGCTGCCGC
 CGCTGCTGCTGGGGGCCTTCCCACCGGCCCGCCGCCCCGAGGCCCCCAAGATGGCGGAC
 AAGGTGGTCCCACGGCAGGTGGCCCCGCTGGGCCGCACTGTGCGGCTGCAGTGCCAGTGGA
 GGGGACCCGCGCCGCTGACCATGTGGACCAAGGATGGCCGCACCATCCACAGCGGCTGGA
 GCCGCTTCCGCGTGCTGCCGCAGGGGCTGAAGGTGAAGCAGGTGGAGCGGGAGGATGCCGGC
 GTGTACGTGTGCAAGGCCACCAACGGCTTCGGCAGCCTGAGCGTCAACTACACCCTCGTCGT
 GCTGGATGACATTAGCCCAGGGAAGGAGAGCCTGGGGCCCCGACAGCTCCTCTGGGGGTCAAG
 AGGACCCCGCCAGCCAGCAGTGGGCACGACCGCGCTTCACACAGCCCTCCAAGATGAGGCGC
 CGGGTGATCGCACGGCCCCGTGGGTAGCTCCGTGCGGCTCAAGTGCCTGGCCAGCGGGCACCC
 TCGGCCCCGACATCACGTGGATGAAGGACGACAGGCCCTTGACGCGCCAGAGGCCGCTGAGC
 CCAGGAAGAAGAAGTGGACACTGAGCCTGAAGAACCTGCGGCCGAGGACAGCGGCAATAC
 ACCTGCCGCGTGTGCAACC GCGCGGGCGCCATCAACGCCACCTACAAGGTGGATGTGATCCA
 GCGGACCCGTTCCAAGCCGCTGCTCACAGGCACGCACCCCGTGAACACGACGGTGGACTTCG
 GGGGACACAGTCTTCCAGTGCAAGGTGCGCAGCGACGTGAAGCCGGTGATCCAGTGGCTG
 AAGCGCGTGGAGTACGGCGCCGAGGGCGCCACAACCTCCACCATCGATGTGGCGGCCAGAA
 GTTTGTGGTGCTGCCACGGGTGACGTGTGTCGCGGCCGACGGCTCCTACCTCAATAAGC
 TGCTCATCGCCGTGCCGCCAGGACGATGCGGGCATGTACATCTGCCTTGGCGCCAACACC
 ATGGGCTACAGCTTCCGCAGCGCCTTCTCACCGTGCTGCCAGACCCAAAACCGCCAGGGCC
 ACCTGTGGCCTCCTCGTCTCGGCCACTAGCCTGCCGTGGCCCGTGGTTCATCGGCATCCCAG
 CCGGCGCTGTCTTCATCTGGGCACCTGCTCCTGTGGCTTTGCCAGGCCAGAGAAAGCCG
 TGCACCCCGCGCCTGCCCTCCCCCTGCCCTGGGCACCGCCCGCCGCGGACGGCCCGGACCCG
 CAGCGGAGACAAGGACCTTCCCTCGTTGGCCGCCCTCAGCGCTGGCCCTGGTGTGGGGCTGT
 GTGAGGAGCATGGGTCTCCGGCAGCCCCCAGCACTTACTGGGCCAGGCCAGTTGCTGGC
 CCTAAGTTGTACCCCAAACCTCTACACAGACATCCACACACACACACACACTCTCACAC
 ACACTCACACGTGGAGGGCAAGGTCCACCAGCACATCCACTATCAGTGCTAGACGGCACCGT
 ATCTGCAGTGGGCACGGGGGGCGGCCAGACAGGCAGACTGGGAGGATGGAGGACGGAGCT
 GCAGACGAAGGCAGGGGACCCATGGCGAGGAGGAATGGCCAGCACCCAGGCAGTCTGTGTG
 TGAGGCATAGCCCCCTGGACACACACACACAGACACACACACTACCTGGATGCATGTATGCAC
 ACACATGCGCGCACACGTGCTCCCTGAAGGCACACGTACGCACACGCACATGCACAGATATG
 CCGCCTGGGCACACAGATAAGCTGCCCAAATGCACGCACACGCACAGACAGATGCCAGACA
 TACAAGGACATGCTGCCTGAACATACACACGCACACCCATGCGCAGATGTGCTGCCTGGACA
 CACACACACACAGGATATGCTGTCTGGACGCACACAGTGCAGATATGGTATCCGGACACA
 CACGTGCACAGATATGCTGCCTGGACACACAGATAATGCTGCCTTGACACACACATGCACGG
 ATATTGCCTGGACACACACACACACACACGCGTGCACAGATATGCTGTCTGGACACGCACAC
 ACATGCAGATATGCTGCCTGGACACACACTTCCAGACACACGTGCACAGGCGAGATATGCT
 GCCTGGACACACGCAGATATGCTGTCTAGTCACACACACACACGCAGACATGCTGTCCGGACAC
 ACACACGCATGCACAGATATGCTGTCCGGACACACACACGCACGCAGATATGCTGCCTGGAC
 ACACACACAGATAATGCTGCCTCAACACTCACACACGTGCAGATATTGCCTGGACACACACA
 TGTGCACAGATATGCTGTCTGGACATGCACACACGTGCAGATATGCTGTCCGGATACACAGT
 CACGCACACATGCAGATATGCTGCCTGGGCACACACTTCCGGACACACATGCACACACAGGT
 GCAGATATGCTGCCTGGACACACACACAGATAATGCTGCCTCAACACTCACACACGTGCAGA
 TATTGCCTGGACACACACATGTGCACAGATATGCTGTCTGGACATGCACACACGTGCAGATA
 TGCTGTCCGGATACACACGCACGCACACATGCAGATATGCTGCCTGGGCACACACTTCCGGA
 CACACATGCACACACAGGTGCAGATATGCTGCCTGGACACACGCAGACTGACGTGCTTTTGG
 GAGGGTGTGCCGTGAAGCCTGCAGTACGTGTGCCGTGAGGCTCATAGTTGATGAGGGACTTT
 CCCTGCTCCACCGTCACTCCCCCAACTCTGCCCGCCTCTGTCCCCGCTCAGTCCCCGCCTC
 CATCCCCGCTCTGTCCCCTGGCCTTGGCGGCTATTTTTGCCACCTGCCTTGGGTGCCCAGG
 AGTCCCCTACTGCTGTGGGCTGGGGTTGGGGGCACAGCAGCCCCAAGCCTGAGAGGCTGGAG
 CCCATGGCTAGTGGCTCATCCCCAGTGCATTCTCCCCCTGACACAGAGAAGGGGCTTGTA
 TTTATATTTAAGAAATGAAGATAATATTAATAATGATGAAGGAAGACTGGGTGAGGGAC
 TGTGGTCTCTCCTGGGGCCCGGGACCCGCTGGTCTTTTACGCCATGCTGATGACCACACCCC
 GTCCAGGCCAGACACCACCCCCACCCCACTGTCGTGGTGGCCCCAGATCTCTGTAATTTTA
 TGTAGAGTTTGAAGCCCCGTATATTTAATTTATTTTGTAAACACAAAA

FIGURE 70

MTPSPLLLLLLLPLLLLGAFPPAAAARGPPKMADKVVPRQVARLGRTVRLQCPVEGDPPPLTM
WTKDGRTIHSGWSRFRVLPQGLKVKQVEREDAGVYVCKATNGFGSLSVNYTLVVLDDISPGK
ESLGPDSSSSGGQEDPASQQWARPRFTQPSKMRRRVIARPVGSSVRLKCVASGHPRPDITWMK
DDQALTRPEAAEPRKKKWTLSLKNLRPEDSGKYTCRVSNRAGAINATYKVDVIQRTRSKPVL
TGTHPVNTTVDFGGTTSFQCKVRSDVKPVIQWLKRVEYGAEGRHNSTIDVGGQKFVVLPTGD
VWSRPDGSYLNKLLITRARQDDAGMYICLGANTMGYSFRSAFLTIVLPDPKPPGPPVASSSSA
TSLPWPVVIGIPAGAVFILGTTTTLLWLCQAQKKPCTPAPAPPLPGHRPPGTARDRSGDKDLPS
LAALSAGPGVGLCEEHGSPAAPQHLLGPGPVAGPKLYPKLYTDIHTHTHTHSHTHSHVEGKV
HQHIHYQC

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 72

MVGTKAWVFSFLVLEVTSVLGRQTMLTQSVRRVQPGKKNPSIFAKPADTLESPGEWTTWFNI
DYPGGKGDYERLDAIRFYYGDRVCARPLRLEARTTDWTPAGSTGQVVHGSPREGFWCLNREQ
RPGQNC SNYTVRFLCPPGSLRRDTERIWS PWS PWSKCSAACGQTGVQTRTRICLAEMVSLCS
EASEEGQHCMGQDCTACDLTCPMGQVNADCDACMCQDFMLHGAVSLPGGAPASGAAIYLLTK
TPKLLTQTDSDGRFRIPGLCPDGKSILKITKVKFAPIVL TMPKTS LKAAT IKA EFVRAETPY
MVMNPETKARRAGQSVSLCCKATGKPRPD KYFWYHNDTLLDPSLYKHESKLVLRKLQQHQAG
EYFCKAQSDAGAVKSKVAQLIVTASDETPCNVPVPE SYLIRLPHDCFQ NATNSFY YDVGRCPV
KTCAGQQDNGIRCRDAVQNC CGISKTEERE IQCSGYTLPTKVAKECSCQRCTETRSIVRGRV
SAADNGEPMRFGHVYMGNSRVSM TGYKGTFTLHVPQDTERLVLT FVDR LQKFVNTTKVLPFN
KKGSAVFHEIKMLRRKEPIT LEAMETNII PLGEVVGEDPMAELEIPSR SFYRQNGEPYIGKV
KASVTFLDPRNISTATAAQTDLNF INDEGDTFPLRTYGMFSVDFRDEVTSEPLNAGKVKVHL
DSTQVKMPEHISTVKLWSLNPDTGLWEEEGDFK FENQRRNKREDRTFLVGNLEIRERRLFNL
DVPESRRCFVKVRAYRSE RFLPSEQIQGVVISVINLEPRTGF LSNPRAWGRFDSVITGPNGA
CVPAFCDDQSPDAYSAYVLASLAGEELQAVESSPKFNPNAIGVPQPYLNKLN YRRTDHEDPR
VKKTAFQISMAKPRPNSAEESNGPIYAFENLRACEEAPPSAAHFRFYQIEGDRYDYNTVPFN
EDDPMSWTE DYLAWWPKPMEFRACYIKVKIVGPLEVNVR SRNMGGTHRRTVGKLYGIRDVRS
TRDRDQPNVSAACLEFKCSGMLYDQDRVDRTL VKVIPQGS CRRASVNPMLHEYLVNHLPLAV
NNDTSEYTMLAPLDPLGHNYGIYTVTDQDPRTAKEIALGRCFDGTSDGSSRIMKSNVGVALT
FNCVERQVGRQSAFQYLQSTPAQSPAAGTVQGRVPSRRQQRASRGGQRQGGVVASLRFP RVA
QQPLIN

FIGURE 73

CTGCAAGTTGTTAACGCCTAACACACAAGTATGTTAGGCTTCCACCAAAGTCCTCAATATACCTGAATACGCAC
 AATATCTTAACTCTTCATATTTGGTTTTGGGATCTGCTTTGAGGTCCCATCTTCATTTAAAAAAAATACAGAG
 ACCTACCTACCCGTACGCATACATACATATGTGTATATATATGTAACTAGACAAAGATCGCAGATCATAAAGC
 AAGCTCTGCTTTAGTTTTCCAAGAAGATTACAAAGAATTTAGAGATGTTTGTCAAGATCCCTGTCGATTTCATG
 CCCTTTGGGTTACGGTGTCTCAGTGATGCAGCCCTACCCTTTGGTTTGGGGACATTATGATTTGTGTAAGACT
 CAGATTTACACGGAAGAGGGAAAAGTTTGGGATTACATGGCCTGCCAGCCGGAATCCACGGACATGACAAAATA
 TCTGAAAGTGAACTCGATCCTCCGGATATTACCTGTGGAGACCCTCCTGAGACGTTCTGTGCAATGGGCAATC
 CCTACATGTGCAATAATGAGTGTGATGCGAGTACCCCTGAGCTGGCACACCCCCCTGAGCTGATGTTTGATTTT
 GAAGGAAGACATCCCTCCACATTTTGGCAGTCTGCCACTTGGGAAGGAGTATCCCAAGCCTCTCCAGGTTAACAT
 CACTCTGTCTTGGAGCAAAACCATTGAGCTAACAGACAACATAGTTATTACCTTTGAATCTGGGCGTCCAGACC
 AAATGATCCTGGAGAAGTCTCTCGATTATGGACGAACATGGCAGCCCTATCAGTATTATGCCACAGACTGCTTA
 GATGCTTTTTCACATGGATCCTAAATCCGTGAAGGATTTATCACAGCATACGGTCTTAGAAATCATTTGCACAGA
 AGAGTACTCAACAGGGTATACAACAAATAGCAAAATAATCCACTTTGAAATCAAAGACAGGTTTCGCGCTTTTGTG
 CTGGACCTCGCCTACGCAATATGGCTTCCCTCTACGGACAGCTGGATACAACCAAGAAACTCAGAGATTTCTTT
 ACAGTCACAGACCTGAGGATAAGGCTGTTAAGACCAGCCGTTGGGGAAATATTTGTAGATGAGCTACACTTGGC
 ACGCTACTTTTACGCGATCTCAGACATAAAGGTGCGAGGAAGGTGCAAGTGTAACTCCATGCCACTGTATGTG
 TGTATGACAACAGCAAATTGACATGCGAATGTGAGCACAACACTACAGGTCCAGACTGTGGGAAATGCAAGAAG
 AATTATCAGGGCCGACCTTGGAGTCCAGGCTCCTATCTCCCCATCCCCAAAGGCACTGCAAATACCTGTATCCC
 CAGTATTTCCAGTATTGGTACGAATGTCTGCGACAACGAGCTCCTGCACTGCCAGAACGGAGGGACGTGCCACA
 ACAACGTGCGCTGCCGTGTGCCCGGCCCATACACGGGCATCCTCTGCGAGAAGCTGCGGTGCCAGGAGGCTGGC
 AGCTGCGGCTCCGACTCTGGCCAGGGCGCGCCCCCGCACGGCACCCAGCGCTGCTGCTGCTGACCACGCTGCT
 GGGAAACCGCCAGCCCCCTGGTGTCTTAGGTGTACCTCCAGCCACACCCGACGGGCCTGTGCCGTGGGGAAGCA
 GACACAACCCAAACATTTGCTACTAACATAGGAAACACACACATACAGACACCCCCACTCAGACAGTGTACAAA
 CTAAGAAGGCCTAACTGAACTAAGCCATATTTATCACCCGTGGACAGCACATCCGAGTCAAGACTGTTAATTTTC
 TGACTCCAGAGGAGTTGGCAGCTGTTGATATTATCACTGCAAATCACATTGCCAGCTGCAGAGCATATTGTGGA
 TTGGAAAGGCTGCGACAGCCCCCAACAGGAAAGACAAAAACAAACAAATCAACCGACCTAAAAACATTGGC
 TACTCTAGCGTGGTGCGCCCTAGTACGACTCC3CCAGTGTGTGGACCAACCAATAGCATTTCTTTGCTGTGAG
 GTGCATTGTGGGCATAAGGAAATCTGTTACAAGCTGCCATATTGGCCTGCTTCCGTCCCTGAATCCCTTCCAAC
 CTGTGCTTTAGTGAACGTTGCTCTGTAAACCTCGTTGGTTGAAAGATTTCTTTGTCTGATGTTAGTGATGCACA
 TGTGTAACAGCCCCCTCTAAAAGCGCAAGCCAGTCATACCCCTGTATATCTTAGCAGCACTGAGTCCAGTGCGA
 GCACACACCCACTATACAAGAGTGGCTATAGGAAAAAAGAAAGTGTATCTATCCTTTTGTATTCAAATGAAGTT
 ATTTTTCTTGAACACTGTAATATGTAGATTTTTTGTATTATTGCCAATTTGTGTTACCAGACAATCTGTTAAT
 GTATCTAATTGGAATCAGCAAAGACTGACATTTTATTTTGTCTCTTTCTGTTTGTGTTTCACTGTGCAGA
 GATTTCTCTGTAAGGGCAACGAACGTGCTGGCATCAAAGAATATCAGTTTACATATATAACAAGTGTAAATAAGA
 TTCCACCAAAGGACATTCTAAATGTTTTCTTGTGTGCTTTAACACTGGAAGATTTAAAGAATAAAAACTCCTGCA
 TAAACGATTTTCAAGGAATTTGTATTGCAATTTCTTAAGATGAAAGGAACAGCCACCAAGCAGTTTCACTCACT
 TTAAGTATTTCTGTGTGGACTGAGTACATTGAGCTGACGAATTTAGTTCCCAGGAAGATGGATTGATGTTCACT
 AGCTTGGACAACCTTCTGCAAAATATGAGACTATTTCCACTTGGGAAAAATTACAACAGCAAAAAAAAAAAAAA
 AAAAAA

FIGURE 74

MYLSRSLSIHALWVTVSSVMQPYPLVWGHYDLCKTQIYTEEGKVWDYMACQPESTDMTKYLK
VKLDPPDITCGDPPETFCAMGNPYMCNNECDASTPELAHPPELMFDFEGRHPSTFWQSATWK
EYPKPLQVNITLSWSKTIELTDNIVITFESGRPDQMILEKSLDYGRTWQPYQYYATDCLDAF
HMDPKSVKDLSQHTVLEIICTEEYSTGYTTNSKIIHFEEKDRFALFAGPRLRNMASLYGOLD
TTKKLRDFFTVTDLRIRLLRPAVGEIFVDELHLARYFYAISDIKVRGRCKCNLHATVCVYDN
SKLTCECEHNTTGPDCGKCKKNYQGRPWSPGSYLPIPKGTANTCIPSISSIGTNVCDNELLH
CQNGGTCHNNVRCLCPAAYTGILCEKLRCEEAGSCGSDSGQGAPPHGTPALLLLTTLLGTAS
PLVF

74/330

FIGURE 75

CCCACGCGTCCGGGTGACCTGGGCCGAGCCCTCCCGGTCCGGCTAAGATTGCTGAGGAGGCGG
CGGGTAGCTGGCAGGCGCCGACTTCCGAAGGCCGCCGTCCGGGCGAGGTGTCCTCATGACTT
CTCTTGTGGACCATGTCCGTGATCTTTTTTGCCTGCGTGGTACGGGTAAGGGATGGACTGCC
CCTCTCAGCCTCTACTGATTTTTTACCACACCCAAGATTTTTTGGAAATGGAGGAGACGGCTCA
AGAGTTTAGCCTTGGCGACTGGCCAGTATCCAGGTGAGGTCTGCAGAAGGTTGTGACTTT
AGTATACATTTTTTCTTCTTTCGGGGACGTGGCCTGCATGGCTATCTGCTCCTGCCAGTGTCC
AGCAGCCATGGCCTTCTGCTTCCTGGAGACCCTGTGGTGGGAATTCACAGCTTCCTATGACA
CTACCTGCATTGGCCTAGCCTCCAGGCCATACGCTTTTTCTTGAGTTTGACAGCATCATTCAG
AAAGTGAAGTGGCATTTTAACTATGTAAGTTCCTCTCAGATGGAGTGCAGCTTGGAATAAT
TCAGGAGGAGCTCAAGTTGCAGCCTCCAGCGGTTCTCACTCTGGAGGACACAGATGTGGCAA
ATGGGGTGATGAATGGTCACACACCGATGCACTTGGAGCCTGCTCCTAATTTCCGAATGGAA
CCAGTGACAGCCCTGGGTATCCTCTCCCTCATTCTCAACATCATGTGTGCTGCCCTGAATCT
CATTCGAGGAGTTCACCTTGCAGAACATTCTTTACAGGATCCAAGGAGCTGGTTCTGCTGGT
TGGACCAAACCTCGTAGGCCAGCCACCCCTGACCCAAATGAGGAGAGCTCTGATTCTCCCAT
CCGGGAGCAGTGATGTCAAACCTTCTGCTGCTGGGGAAATCTCATCAGCAGGGAGCCTGTGGA
AAAGGGCATGTCAGTGAAATCTGGGAATGGCTGGATTTCGGAACATCTGCCCATGTGTATTG
ATGGCAGAGCTGTTGCCCACAAGCGCCTTTTATTTAGGGTAAAATTAACAAATCCATTCTAT
TCCTCTGACCCATGCTTAGTACATATGACCTTTAACCCTTACATTTATATGATTCTGGGGTT
GCTTCAGAAGTGTTATTTTCATGAATCATTCATATGATTTGATCCCCCAGGATTCTATTTTGT
TTAATGGGCTTTTCTACTAAAAGCATAAAATACTGAGGCTGATTTAGTCAGGGCAAAACCAT
TTACTTTACATATTCGTTTTCAATACTTGCTGTTTCATGTTACACAAGCTTCTTACGGTTTTT
TTGTAACAATAAATATTTTGAGTAAATAATGGGTACATTTTAACAACTCAGTAGTACAACC
TAACTTGTATAAAAGTGTGTAAAAATGTATAGCCATTTATATCCTATGTATAAATTAATG
AGGTGGCTTCAGAAATGGCAGAATAAATCTAAAGTGTTTATTAATAAAAAAAAAAAAAAAAAA
AAAAG

76/330

FIGURE 76

MSVIFFACVVRVRDGLPLSASTDFYHTQDFLEWRRRLKSLALRLAQYPGRGSAEGCDFS IHF
SSFGDVACMAICSCQCPAAMAFCFLETLWWEFTASYDTTCIGLASRPYAFLEFDSIIQKVW
HFNYVSSSQMECSLEKIQEELKLQPPAVLTLEDTDVANGVMNGHTPMHLEPAPNFRMEPVTA
LGILSLILNIMCAALNLIRGVHLAEHSLQDPRSWFCWLDQTS

MSVIFFACVVRVRDGLPLSASTDFYHTQDFLEWRRRLKSLALRLAQYPGRGSAEGCDFS IHF
SSFGDVACMAICSCQCPAAMAFCFLETLWWEFTASYDTTCIGLASRPYAFLEFDSIIQKVW
HFNYVSSSQMECSLEKIQEELKLQPPAVLTLEDTDVANGVMNGHTPMHLEPAPNFRMEPVTA
LGILSLILNIMCAALNLIRGVHLAEHSLQDPRSWFCWLDQTS

77/330

FIGURE 77

TGCTTCCTGGAGACCCTGTGGTGGGAATTCACAGCTTCNTATGACACTACCTGCATTGGCNT
AGCCTCCAGGCCATACGCTTTTCTTGAGTTTGACAGCATCATTCAGAAAGTGAAGTGGCATT
TTAACTATGTAAGTTCCTNTCAGATGGAGTGCAGCTTGAAAAAATTCAGGAGGAGCTCAAG
TTGCAGCCTCCAGCGGTTCTCANTATGGAGGACACAGATGTGGCAAATGGGGT

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 78

CTCAGCGGCGCTTCCTCGTAGCGAGCCTAGTGGCGGGTGTTCGATTGAAACGTGAGCGCGA
 CCCGACCTTAAAGAGTGGGGAGCAAAGGGAGGACAGAGCCCTTTAAAACGAGGCGGGTGGTG
 CCTGCCCCTTTAAAGGGCGGGGCGTCCGGACGACTGTATCTGAGCCCCAGACTGCCCCGAGTT
 TCTGTTCGAGGCTGCGAGGAAAGGCCCCCTAGGCTGGGTCTGGGTGCTTGGCGGGCGGCGGCTT
 CCTCCCCGCTCGTCTCCTCCCCGGGCCCAGAGGCACCTCGGCTTCAGTCATGCTGAGCAGAGTA
TGGAAGCACCTGACTACGAAGTGCTATCCGTGCGAGAACAGCTATTCCACGAGAGGATCCGC
 GAGTGTATTATATCAACACTTCTGTTTGCAACACTGTACATCCTCTGCCACATCTTCCTGAC
 CCGCTTCAAGAAGCCTGCTGAGTTCACCACAGTGGATGATGAAGATGCCACCGTCAACAAGA
 TTGCGCTCGAGCTGTGCACCTTTACCCTGGCAATTGCCCTGGGTGCTGTCTGCTCCTGCCC
 TTCTCCATCATCAGCAATGAGGTGCTGCTCTCCCTGCCTCGGAACACTACTACATCCAGTGGCT
 CAACGGCTCCCTCATCCATGGCCTCTGGAACCTTGTTTTCTCTTCCCCAACCTGTCCCTCA
 TCTTCCTCATGCCCTTTGCATATTTCTTCACTGAGTCTGAGGGCTTTGCTGGCTCCAGAAAG
 GGTGTCTGTTGGGCGGGTCTATGAGACAGTGGTGATGTTGATGCTCCTCACTCTGCTGGTGCT
 AGGTATGGTGTGGGTGGCATCAGCCATTGTGGACAAGAACAAGGCCAACAGAGAGTCACTCT
 ATGACTTTTGGGAGTACTATCTCCCCTACCTCTACTCATGCATCTCCTTCCTTGGGGTTCTG
 CTGCTCCTGGTGTGTACTCCACTGGGTCTCGCCCGCATGTTCTCCGTCACTGGGAAGCTGCT
 AGTCAAGCCCCGGCTGCTGGAAGACCTGGAGGAGCAGCTGTACTGCTCAGCCTTTGAGGAGG
 CAGCCCTGACCCGCAGGATCTGTAATCCTACTTCTGCTGGCTGCCTTTAGACATGGAGCTG
 CTACACAGACAGGTCCTGGCTCTGCAGACACAGAGGGTCCTGCTGGAGAAGAGGCGGAAGGC
 TTCAGCCTGGCAACGGAACCTGGGCTACCCCTGGCTATGCTGTGCTTGTGCTGGTGCTGACGG
 GCCTGTCTGTGCTCATTGTGGCCATCCACATCCTGGAGCTGCTCATCGATGAGGCTGCCATG
 CCCCAGGGCATGCAGGGTACCTCCTTAGGCCAGGTCTCCTTCTCCAAGCTGGGCTCCTTTGG
 TGCCGTCAATTCAGGTTGTACTCATCTTTTACCTAATGGTGTCTCAGTTGTGGGCTTCTATA
 GCTCTCCACTCTTCCGGAGCCTGCGGGCCAGATGGCACGACACTGCCATGACGCAGATAATT
 GGGAACTGTGTCTGTCTCCTGGTCCTAAGCTCAGCACTTCCTGTCTTCTCTCGAACCTGGG
 GCTCACTCGCTTTGACCTGCTGGGTGACTTTGGACGCTTCAACTGGCTGGGCAATTTCTACA
 TTGTGTTCTCTACAACGCAGCCTTTGCAGGCCTCACCACACTCTGTCTGGTGAAGACCTTC
 ACTGCAGCTGTGCGGGCAGAGCTGATCCGGGCCTTTGGGCTGGACAGACTGCCGCTGCCCCGT
 CTCCGGTTTCCCCCAGGCATCTAGGAAGACCCAGCACCAGTGAACCTCCAGCTGGGGGTGGGA
 AGGAAAAAACTGGACACTGCCATCTGCTGCCTAGGCCTGGAGGGAAGCCCAAGGCTACTTGG
 ACCTCAGGACCTGGAATCTGAGAGGGTGGGTGGCAGAGGGGAGCAGAGCCATCTGCACTATT
 GCATAATCTGAGCCAGAGTTTGGGACCAGGACCTCCTGCTTTTCCATACTTAACTGTGGCCT
 CAGCATGGGGTAGGGCTGGGTGACTGGGTCTAGCCCCTGATCCCAAATCTGTTTACACATCA
 ATCTGCCTCACTGCTGTTCTGGGCCATCCCCATAGCCATGTTTACATGATTTGATGTGCAAT
 AGGGTGGGGTAGGGGCAGGGAAGGACTGGGCCAGGGCAGGCTCGGGAGATAGATTGTCTCC
 CTTGCCTCTGGCCCAGCAGAGCCTAAGCACTGTGCTATCCTGGAGGGGCTTTGGACCACCTG
 AAAGACCAAGGGGATAGGGAGGAGGAGGCTTCAGCCATCAGCAATAAAGTTGATCCCAGGGA
 AAAAAA

79/330

FIGURE 79

MEAPDYEVLSVREQLFHERIRECIIISTLLFATLYILCHIFLTRFKKPAEFTTVDDDEDATVNK
IALELCTFTLAIALGAVLLLPFSIISNEVLLSLPRNYYIQWLNGSLIHGLWNLVFLFPNLSL
IFLMPFAYFFTESEGFAGSRKGVLRVYETVVMLMLLTLLVLGMVWVASAIVDKNKANRESL
YDFWEYYLPYLYSCISFLGVLLLLVCTPLGLARMFSVTGKLLVKPRLLEDLEEQLYCSAFEE
AALTRRICNPTSCWLPLDMELLHRQVLALQTQRVLLEKRRKASAWQRNLGYPLAMLCLLVLT
GLSVLIVAIHILELLIDEAAMPQGTSLGQVSFSKLGSGFAVIQVVLIIFYLMVSSVVGFY
SSPLFRSLRPRWHDAMTQIIIGNCVCLLVSSALPVFSRTLGLTRFDLLGDFGRFNWLGNFY
IVFLYNAAFAGLTTLCLVKTFATAAVRAELIRAFGLDRLPLPVSGFPQASRKTQHQ

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

80/330

FIGURE 80

GGCTGCCGAGGGAAGGCCCTTGGGTGGTCTTGGTTGCTTGGCGGCGGCGGNTTCNTCCCC
GCTCGTCCTCCCCGGGCCCAGAGGCACCTCGGCTTCAGTCATGCTGAGCAGAGTATGGAAGC
ACCTGACTACGAAGTGCTATCCGTGCGAGAACAGCTATTCCACGAGAGGATCCGCGAGTGTA
TTATATCAACACTTCTGTTTGCAACACTGTACATCCTCTGCCACATCTTCCTGACCCGCTTC
AAGAAGCCTGCTGAGTTCACCACAGTGGATGATGAAGATGCCACCG

Figure 80

FIGURE 81

GACCGACCTTAAAGAGTGGGAGCAAAGGGAGGACAGAGCCTTTTAAAACGAGGCGGTGGTGC
CTGCCCTTTAAGGGCGGGGCGTCCGGACGACTGTATCTGAGCCCCAGACTGCCCCGAGTTTC
TGTCGCAGGCTGCGAGGAAAGGCCCTAGGCTGGGTCTGGTGCTTGGCGGCGGCGGCTTCCT
CCCCGTTGTCNTCCCCGGGCCCAGAGGCACCTCGGCTTCAGTCATGCTGAGCAGAGTATGGA
AGCACCTGACTACGAAGTGCTATCCGTGCGAGAACAGCTATTCCACGAGAGGATCCGCGAGT
GTATTATATCAACACTTCTGTTTGCAACACTGTACATCNTCTGCCACATCTTCCTGACCCGC
TTCAAGAAGCCTGCTGAGTTCACCACAGTGGATGATGAAGATGCCACCGTCAACAAGATTGC
GCTCGAGCTGTGCACCTTTACCCTGGCAATTGCCCTGGGTGCTGTCTCTGCTCCTGCCCTTCT
CCATCATCAGCAATGAGGTGCTGCACTCCC

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 82

GATGTGCTCCTTGGAGCTGGTGTGCAGTGTCTGACTGTAAGATCAAGTCCAAACCTGTTTT
GGAATTGAGGAACTTCTCTTTTGATCTCAGCCCTTGGTGGTCCAGGTCTTC**ATG**CTGCTGT
GGGTGATATTACTGGTCCTGGCTCCTGTCAGTGGACAGTTTGCAAGGACACCCAGGCCCATTT
ATTTTCCTCCAGCCTCCATGGACCACAGTCTTCCAAGGAGAGAGAGTGACCCTCACTTGCAA
GGGATTTTCGCTTCTACTCACCACAGAAAACAAAATGGTACCATCGGTACCTTGGGAAAGAAA
TACTAAGAGAAACCCAGACAATATCCTTGAGGTTCAGGAATCTGGAGAGTACAGATGCCAG
GCCAGGGCTCCCCTCTCAGTAGCCCTGTGCACTTGGATTTTTCTTCAGAGATGGGATTTCC
TCATGCTGCCCAGGCTAATGTTGAACTCCTGGGCTCAAGTGATCTGCTCACCC**TAG**GCCTCTC
AAAGCGCTGGGATTACAGCTTCGCTGATCCTGCAAGCTCCACTTTCTGTGTTTGAAGGAGAC
TCTGTGGTTCTGAGGTGCCGGGCAAAGGCGGAAGTAACACTGAATAATACTATTTACAAGAA
TGATAATGTCCTGGCATTCCCTTAATAAAAGAACTGACTTCCAAAAAAAAAAAAAAAAAAAAA
AAA

FIGURE 84

CAGAAGAGGGGGCTAGCTAGCTGTCTCTGCGGACCAGGGAGACCCCCGCGCCCCCCCCGGTGT
GAGGCGGCCTCACAGGGCCGGGTGGGCTGGCGAGCCGACGCGGCGGCGGAGGAGGCTGTGAG
GAGTGTGTGGAACAGGACCCGGGACAGAGGAACCATGGCTCCGCAGAACCTGAGCACCTTTT
GCCTGTTGCTGCTATACCTCATCGGGGCGGTGATTGCCGGACGAGATTTCTATAAGATCTTG
GGGTGCCTCGAAGTGCCTCTATAAAGGATATTAAAAAGGCCTATAGGAACTAGCCCTGCA
GCTTCATCCCGACCGGAACCCTGATGATCCACAAGCCCAGGAGAAATTCAGGATCTGGGTG
CTGCTTATGAGGTTCTGTGAGATAGTGAGAAACGGAAACAGTACGATACTTATGGTGAAGAA
GGATTAAAAGATGGTCATCAGAGCTCCCATGGAGACATTTTTTTCACACTTCTTTGGGGATTT
TGGTTTCATGTTTGGAGGAACCCCTCGTCAGCAAGACAGAAATATTCCAAGAGGAAGTGATA
TTATTGTAGATCTAGAAGTCACTTTGGAAGAAGTATATGCAGGAAATTTTGTGGAAGTAGTT
AGAAACAAACCTGTGGCAAGGCAGGCTCCTGGCAAACGGAAGTGCAATTGTGCGCAAGAGAT
GCGGACCACCCAGCTGGGCCCTGGGCGCTTCCAAATGACCCAGGAGGTGGTCTGCGACGAAT
GCCCTAATGTCAAACAGTGAATGAAGAACGAACGCTGGAAGTAGAAATAGAGCCTGGGGTG
AGAGACGGCATGGAGTACCCCTTTATTGGAGAAGGTGAGCCTCACGTGGATGGGGAGCCTGG
AGATTTACGGTTCCGAATCAAAGTTGTCAAGCACCCAATATTTGAAAGGAGAGGAGATGATT
TGTACACAAATGTGACAATCTCATTAGTTGAGTCACTGGTTGGCTTTGAGATGGATATTACT
CACTTGGATGGTCACAAGGTACATATTTCCCGGGATAAGATCACCAGGCCAGGAGCGAAGCT
ATGGAAGAAAGGGGAAGGGCTCCCCAACTTTGACAACAACAATATCAAGGGCTCTTTGATAA
TCACTTTTGATGTGGATTTTCCAAAAGAACAGTTAACAGAGGAAGCGAGAGAAGGTATCAAA
CAGCTACTGAAACAAGGGTCAGTGCAGAAGGTATACAATGGACTGCAAGGATATTGAGAGTG
AATAAAATTGGACTTTGTTTAAAATAAGTGAATAAGCGATATTTATTATCTGCAAGGTTTTT
TTGTGTGTGTTTTTGTTTTTATTTTCAATATGCAAGTTAGGCTTAATTTTTTTATCTAATGA
TCATCATGAAATGAATAAGAGGGCTTAAGAATTTGTCCATTTGCATTCGGAAAAGAATGACC
AGCAAAAGGTTTACTAATACCTCTCCCTTTGGGGATTTAATGTCTGGTGCTGCCGCTGAGT
TTCAAGAATTAAAGCTGCAAGAGGACTCCAGGAGCAAAAGAAACACAATATAGAGGGTTGGA
GTTGTTAGCAATTTCAATCAAAATGCCAACTGGAGAAGTCTGTTTTTAAATACATTTTGTTG
TTATTTT

FIGURE 85

MAPQNLSTFCLLLLYLIGAVIAGRDFYKILGVPRSASIKDIKKAYRKLALQLHPDRNPDDPQ
AQEKFDLGAAYEVLSDSEKQYDQTYGEEGLKDGHQSSHGDI FSHFFGDFGFMFGGTTPRQQ
DRNIPRGSDIIVDLEVTLEEVYAGNFVEVVRNKPVARQAPGKRKCNCRQEMRTTQLGPGRFQ
MTQEVVCDECPNVKLVNEERTLEVEIEPGVRDGMETPFGEGEHPVDGEPGDLRFRIKVVKH
PIFERRGDDLYTNVTISLVESLVGFEMDITHLDGHKVHISRDKITRPGAKLWKKGEGLPNFD
NNNIKGSIIITFDVDFPKEQLTEEAREGIKQLLKQGSVQKVYNGLOGY

Important features:**Signal peptide:**

amino acids 1-22

Cell attachment sequence.

amino acids 254-257

Nt-dnaJ domain signature.

amino acids 67-87

Homologous region to Nt-dnaJ domain proteins.

amino acids 26-58

N-glycosylation site.

amino acids 5-9, 261-265

Tyrosine kinase phosphorylation site.

amino acids 253-260

N-myristoylation site.

amino acids 18-24, 31-37, 93-99, 215-221

Amidation site.

amino acids 164-168

86/330

FIGURE 86

TGGGACCAGGGAACCCCGGGCCCCCGGTGGAGNGCCTAACAGGCCGGTGGNTGCGACCGAA
GCGGCCGGCGGAGGAGGTTTTGAGGATTTTGGAAACAGGACCCGGACAGAGGAACCATGGTT
CCGCAGAACNTGAGCACNTTTTGCCTGTTGNTGNTATACTTCATCGGGGCGGTGATTGCCGG
ACGAGATTTNTATAAGATTTTGGGGTGCCTNGAAGTGCCTTNTATAAAGGATATTAAAAAGG
CCTATAGGAAACTAGCCCTGCAGNTTATCCCGACCGGAACCCTGATGATCCACAAGCCCAG
GAGAAATTCAGGATTTGGGTGCTGCTTATGAGGTTNTGTCAGATAGTGAGAAACGGAAACA
GTACGATAATTATGGTGAAGAAGGATTAAAAGATGGTNATCAGAGCTCCCATGGAGACATTT
TTTCACACTTNTTTGGGGATTTTGGTTTCATGTTTGGAGGAACCCCTNGTCAGCAAGACAGA
AATATTCCAAGAG

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 87

GGCACGAGGCGGCGGGGCAGTCGCGGGATGCGCCCCGGGAGCCACAGCCTGAGGCCCTCAGGT
CTCTGCAGGTGTCGTGGAGGAACCTAGCACCTGCCATCCTCTTCCCCAATTTGCCACTTCCA
GCAGCTTTAGCCCATGAGGAGGATGTGACCGGGACTGAGTCAGGAGCCCTCTGGAAGC**ATGG**
AGACTGTGGTGATTGTTGCCATAGGTGTGCTGGCCACCATCTTTCTGGCTTCGTTTGCAGCC
TTGGTGCTGGTTTGCAGGCAGCGCTACTGCCGGCCGCGAGACCTGCTGCAGCGCTATGATTC
TAAGCCCATTGTGGACCTCATTGGTGCCATGGAGACCCAGTCTGAGCCCTCTGAGTTAGAAC
TGGACGATGTCGTTATCACCAACCCCCACATTGAGGCCATTCTGGAGAATGAAGACTGGATC
GAAGATGCCTCGGGTCTCATGTCCCACTGCATTGCCATCTTGAAGATTTGTCACACTCTGAC
AGAGAAGCTTGTTGCCATGACAATGGGCTCTGGGGCCAAGATGAAGACTTCAGCCAGTGTCA
GCGACATCATTGTGGTGGCCAAGCGGATCAGCCCCAGGGTGGATGATGTTGTGAAGTCGATG
TACCCTCCGTTGGACCCCCAACTCCTGGACGCACGGACGACTGCCCTGCTCCTGTCTGTCAG
TCACCTGGTGCTGGTGACAAGGAATGCCTGCCATCTGACGGGAGGCCTGGACTGGATTGACC
AGTCTCTGTCTGGCTGCTGAGGAGCATTGGAAGTCCTTCGAGAAGCAGCCCTAGCTTCTGAG
CCAGATAAAGGCCTCCCAGGCCCTGAAGGCTTCCTGCAGGAGCAGTCTGCAATT**TAG**GCCT
ACAGGCCAGCAGCTAGCCATGAAGGCCCTGCCGCCATCCCTGGATGGCTCAGCTTAGCCTT
CTACTTTTTCTATAGAGTTAGTTGTTCTCCACGGCTGGAGAGTTCAGCTGTGTGTGCATAG
TAAAGCAGGAGATCCCCGTCAGTTTATGCCTCTTTTGCAAGAGGAGTATTGAAAA
CTGGTGGACTGTCAGCTTTATTTAGCTCACCTAGTGTTTTCAAGAAAATTGAGCCACCGTCT
AAGAAATCAAGAGGTTTCACATTAAAATTAGAATTTCTGGCCTCTCTCGATCGGTCAGAATG
TGTGGCAATTCTGATCTGCATTTTCAGAAGAGGACAATCAATTGAAACTAAGTAGGGGTTTC
TTCTTTTGGCAAGACTTGTAAGTCTCTCACCTGGCCTGTTTCATTTATTTGTATTATCTGCCT
GGTCCCTGAGGCGTCTGGGTCTCTCCTCTCCCTTGCAAGTTTGGGTTTGAAGCTGAGGAACT
ACAAAGTTGATGATTTCTTTTTTATCTTTATGCCTGCAATTTTACCTAGCTACCACTAGGTG
GATAGTAAATTTATACTTATGTTTCCCTCAAAAAAAAAAAAAAA

88/330

FIGURE 88

METVVIVAIGVLATIFLASFAALVLVCRQRYCRPRDLLQRYDSKPIVDLIGAMETQSEPSEL
ELDDVVITNPHIEAILENEDWIEDASGLMSHCIAILKICHTLTEKLVAMTMGSGAKMKT SAS
VSDIIVVAKRISPRVDDVVKSMYPPLDPKLLDARTTALLSVSHLVLVTRNACHLTGGLDWI
DQSLSAEEHLEVLREAALASEPDKGLPGPEGFLQEQSAI

88/330

FIGURE 89

GCTTCATTTCTCCCGACTCAGCTTCCCACCCTGGGCTTCCGAGGTGCTTTCGCCGCTGTCC
CCACCACTGCAGCCATGATCTCCTTAACGGACACGCAGAAAATTGGAATGGGATTAACAGGA
TTTGGAGTGTTTTCTCTGTTCTTTGGAATGATTCTCTTTTTTGACAAAGCACTACTGGCTAT
TGGAATGTTTTATTTGTAGCCGGCTTGGCTTTTGTAATTGGTTTAGAAAGAACATTTCAGAT
TCTTCTTCCAAAAACATAAAATGAAAGCTACAGGTTTTTTTTCTGGGTGGTGTATTTGTAGTC
CTTATTGGTTGGCCTTTGATAGGCATGATCTTCGAAATTTATGGATTTTTTCTCTTGTTTTCAG
GGGCTTCTTCTCTGTCGTTGTTGGCTTTATTAGAAGAGTGCCAGTCCTTGGATCCCTCCTAAAT
TTACCTGGAATTAGATCATTTGTAGATAAAGTTGGAGAAAGCAACAATATGGTATAACAACA
AGTGAATTTGAAGACTCATTTAAAATATTGTGTTATTTATAAAGTCATTTGAAGAATATTCA
GCACAAAATTAAATTACATGAAATAGCTTGTAATGTTCTTTACAGGAGTTTAAAACGTATAG
CCTACAAAGTACCAGCAGCAAATTAGCAAAGAAGCAGTGAAAACAGGCTTCTACTCAAGTGA
ACTAAGAAGAAGTCAGCAAGCAAAGTCTGAGAGAGGTGAAATCCATGTTAATGATGCTTAAGAA
ACTCTTGAAGGCTATTTGTGTTGTTTTTCCACAATGTGCGAAACTCAGCCATCCTTAGAGAA
CTGTGGTGCCTGTTTCTTTTCTTTTTATTTTGAAGGCTCAGGAGCATCCATAGGCATTTGCT
TTTTAGAAGTGTCCACTGCAATGGCAAAAATATTTCCAGTTGCACTGTATCTCTGGAAGTGA
TGCATGAATTCGATTGGATTGTGTCATTTTAAAGTATTAAAACCAAGGAAACCCCAATTTTG
ATGTATGGATTACTTTTTTTTTGNGCNCAGGGCC

90/330

FIGURE 90

MISLTDQTQKIGMGLTGFGVFLLFFGMILFFDKALLAIGNVLFVAGLAFVIGLERTFRFFFQK
HKMKATGFFLGGVFVVLIGWPLIGMIFEIYGFFLLFRGFFPVVVGFI RRVPVLGSLLNLPGI
RSFVDKVGESNNMV

Important features:

Transmembrane domains:

amino acids 12-30 (typeII), 33-52, 69-89 and 93-109

N-myristoylation sites.

amino acids 11-16, 51-56 and 116-121

Aminoacyl-transfer RNA synthetases class-II protein.

amino acids 49-59

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 91

GAAGACGTGGCGGCTCTCGCCTGGGCTGTTTCCCGGCTTCATTTCTCCCGACTCAGCTTCCC
ACCNTGGGCTTTCCGAGGTGCTTTCGCCGCTGTCCCCACCACTGCAGCCATGATCTCCTTAA
CGGACACGCAGAAAATTGGAATGGGATTAACCGGATTTGGAGTGTTTTCTGTTCTTTGGA
ATGATTCTCTTTTTTGACAAAGCACTACTGGCTATTGGAAATGTTTTATTTGTAGCCGGCTT
GGCTTTTGTAATTGGTTTAGAAAGAACATTTCAGATTCTTCTTCCAAAACATAAAATGAAAG
CTACAGGTTTTTTTCTGGGTGGTGTATTTGTAGTCCTTATTGGTTGGCCTTTGATAGGCATG
ATCTTCGAAATTTATGGATTTTTTCTCTTGTTT

91/330

FIGURE 92

GGCACGAGGCTGAACCCAGCCGGCTCCATCTCAGCTTCTGGTTTCTAAGTCCATGTGCCAAA
GGCTGCCAGGAAGGAGACGCCTTCCTGAGTCCTGGATCTTTCTTCCTTCTGGAAATCTTTGA
CTGTGGGTAGTTATTTATTTCTGAATAAGAGCGTCCACGCATC**ATG**GACCTCGCGGGACTGC
TGAAGTCTCAGTTCCTGTGCCACCTGGTCTTCTGCTACGTCTTTATTGCCTCAGGGCTAATC
ATCAACACCATTCAGCTCTTCACTCTCCTCCTCTGGCCCATTAACAAGCAGCTCTTCCGGAA
GATCAACTGCAGACTGTCCTATTGCATCTCAAGCCAGCTGGTGATGCTGCTGGAGTGGTGGT
CGGGCACGGAATGCACCATCTTCACGGACCCGCGCGCCTACCTCAAGTATGGGAAGGAAAAT
GCCATCGTGGTTCTCAACCACAAGTTTGAAATTGACTTTCTGTGTGGCTGGAGCCTGTCCGA
ACGCTTTGGGCTGTTAGGGGGCTCCAAGGTCCTGGCCAAGAAAGAGCTGGCCTATGTCCCAA
TTATCGGCTGGATGTGGTACTTCACCGAGATGGTCTTCTGTTTCGCGCAAGTGGGAGCAGGAT
CGCAAGACGGTTGCCACCAGTTTGCAGCACCTCCGGGACTACCCCGAGAAGTATTTTTTCT
GATTCACCTGTGAGGGCACACGGTTCACGGAGAAGAAGCATGAGATCAGCATGCAGGTGGCCC
GGGCAAGGGGCTGCCTCGCCTCAAGCATCACCTGTTGCCACGAACCAAGGGCTTCGCCATC
ACCGTGAGGAGCTTGAGAAATGTAGTTTCAGCTGTATATGACTGTACACTCAATTTTCAGAAA
TAATGAAAATCCAACACTGCTGGGAGTCCTAAACGGAAAGAAATACCATGCAGATTTGTATG
TTAGGAGGATCCCACTGGAAGACATCCCTGAAGACGATGACGAGTGCTCGGCCTGGCTGCAC
AAGCTCTACCAGGAGAAGGATGCCTTTTCAGGAGGAGTACTACAGGACGGGCACCTTCCCAGA
GACGCCCATGGTGCCCCCCCCGGCGGCCCTGGACCTCGTGAAGTGGCTGTTTTGGGCCTCGC
TGGTGCTCTACCTTTCTTCCAGTTCCTGGTCAGCATGATCAGGAGCGGGTCTTCCCTGACG
CTGGCCAGCTTCATCCTCGTCTTCTTTGTGGCCTCCGTGGGAGTTCGATGGATGATTGGTGT
GACGGAAATTGACAAGGGCTCTGCCTACGGCAACTCTGACAGCAAGCAGAACTGAATGACT**T**
GACTCAGGGAGGTGTCACCATCCGAAGGGAACCTTGGGGAAGTGGTGGCCTCTGCATATCCT
CCTTAGTGGGACACGGTGACAAAGGCTGGGTGAGCCCTGCTGGGCACGGCGGAAGTCACGA
CCTCTCCAGCCAGGGAGTCTGGTCTCAAGGCCGATGGGGAGGAAGATGTTTTGTAATCTTT
TTTTCCCCATGTGCTTTAGTGGGCTTTGGTTTTCTTTTTGTGCGAGTGTGTGTGAGAATGGC
TGTGTGGTGAGTGTGAACCTTTGTTCTGTGATCATAGAAAGGGTATTTTAGGCTGCAGGGGAG
GGCAGGGCTGGGGACCGAAGGGGACAAGTTCCCCTTTCATCCTTTGGTGCTGAGTTTTCTGT
AACCCTTGGTTGCCAGAGATAAAGTGAAAAGTGCTTTAGGTGAGATGACTAAATTATGCCTC
CAAGAAAAAAAATTAAAGTGCTTTTCTGGGTCAAAAAAAAAAAAA

FIGURE 93

MDLAGLLKSQFLCHLVFCYVFIA SGLIINTIQLF TLLLWPINKQLFRKINCRLSYC ISSQLV
MLLEWWSGTECTIFTDPRAYLKYGKENAIVVLN HKFEIDFLCGWSLSERFGLLGGSKVLAKK
ELAYVPIIGWMMWYFTEMVFCSRKWEQDRKTVATSLQH LRDYPEKYFFLIHCEGTRFTEKKHE
ISMQVARAKGLPRLKHHLLPRTKGFAITVRSLRNVVSAVYDCTLNFRNNENPTLLGVLNGKK
YHADLYVRRIPLEDIPEDDDECSAWLHKLYQEKDAFQEEYYRTGTFPETPMVPPRRPWTLVN
WLFWASLVLYPFFQFLVSMIRSGSSLTLASFILVFFVASVGVRWMIGVTEIDKGSAYGNSDS
KQKLND

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 94

CTGAGGCGGCGGTAGC**ATG**GAGGGGGAGAGTACGTCGGCGGTGCTCTCGGGCTTTGTGCTCG
GCGCACTCGCTTTCCAGCACCTCAACACGGACTCGGACACGGAAGGTTTTCTTCTTGGGGAA
GTAAAAGGTGAAGCCAAGAACAGCATTACTGATTCCCAAATGGATGATGTTGAAGTTGTTTA
TACAATTGACATTCAGAAATATATTCCATGCTATCAGCTTTTTAGCTTTTATAATTCTTCAG
GCGAAGTAAATGAGCAAGCACTGAAGAAAATATTATCAAATGTCAAAAAGAATGTGGTAGGT
TGGTACAAATTCGTCGTCATTTCAGATCAGATCATGACGTTTAGAGAGAGGCTGCTTCACAA
AACTTGCAGGAGCATTTTTCAAACCAAGACCTTGTTTTTCTGCTATTAACACCAAGTATAA
TAACAGAAAGCTGCTCTACTCATCGACTGGAACATTCTTATATAAACCTCAAAAAGGACTT
TTTCACAGGGTACCTTTAGTGTTGCCAATCTGGGCATGTCTGAACAACTGGGTATATAAAC
TGTATCAGGTTCCCTGTATGTCCACTGGTTTTAGCCGAGCAGTACAAACACACAGCTCTAAAT
TTTTTGAAGAAGATGGATCCTTAAAGGAGGTACATAAGATAAATGAAATGTATGCTTCATTA
CAAGAGGAATTAAAGAGTATATGCAAAAAGTGGAAGACAGTGAACAAGCAGTAGATAAACT
AGTAAAGGATGTAAACAGATTAAACGAGAAATTGAGAAAAGGAGAGGAGCACAGATTCAGG
CAGCAAGAGAGAAGAACATCCAAAAGACCTCAGGAGAACATTTTTCTTTGTCAGGCATTA
CGGACCTTTTTTCCAAATTCTGAATTTCTTCATTCATGTGTTATGTCTTTAAAAAATAGACA
TGTTTCTAAAAGTAGCTGTAACCTACAACCACCTCTCGATGTAGTAGACAATCTGACCTTAA
TGGTAGAACACACTGACATTCCTGAAGCTAGTCCAGCTAGTACACCACAAATCATTAAAGCAT
AAAGCCTTAGACTTAGATGACAGATGGCAATTCAAGAGATCTCGGTTGTTAGATACACAAGA
CAAACGATCTAAAGCAAATACTGGTAGTAGTAACCAAGATAAAGCATCCAAAATGAGCAGCC
CAGAAACAGATGAAGAAATTGAAAAGATGAAGGGTTTTGGTGAATATTCACGGTCTCCTACA
TTTT**TGA**TCCTTTTAACCTTACAAGGAGATTTTTTTATTTGGCTGATGGGTAAAGCCAAACAT
TTCTATTGTTTTTACTATGTTGAGCTACTTGCAGTAAGTTCATTTGTTTTTACTATGTTTAC
CTGTTTGCAGTAATACACAGATAACTCTTAGTGCATTTACTTCACAAAGTACTTTTTCAAAC
ATCAGATGCTTTTATTTCCAAACCTTTTTTTTACCTTTCACTAAGTTGTTGAGGGGAAGGCT
TACACAGACACATTCTTTAGAATTGGAAAAGTGAGACCAGGCACAGTGGCTCACACCTGTAA
TCCCAGCACTTAGGGAAGACAAGTCAGGAGGATTGATTGAAGCTAGGAGTTAGAGACCAGCC
TGGGCAACGTATTGAGACCATGTCTATTAAAAAATAAAATGGAAAAGCAAGAATAGCCTTAT
TTTCAAAATATGGAAAGAAATTTATATGAAAATTTATCTGAGTCATTAAAATTCTCCTTAAG
TGATACTTTTTTTAGAAGTACATTATGGCTAGAGTTGCCAGATAAAATGCTGGATATCATGCA
ATAAATTTGCAAAACATCATCTAAAATTTAAAAAAAAAAAAAAAAAAAAA

95/330

FIGURE 95

MEGESTSAVLSGFVLGALAFQHLNTDSDTEGFLLGEVKGEAKNSITDSQMDDVEVVYTIDIQ
KYIPCYQLFSFYNSSGEVNEQALKKILSNVKKNVVGWYKFRRHSDQIMTFRERLLHKNLQEH
FSNQDLVFLLLTPSIITESCSTHRLEHSLYKPQKGLFHRVPLVVANLGMSEQLGYKTVSGSC
MSTGFSRAVQTHSSKFFFEEDGSLKEVHKINEMYASLQEELKSICKKVEDSEQAVDKLVKDVN
RLKREIEKRRGAQIQAAREKNIQKDPQENIFLCQALRTFFPNSEFLHSCVMSLKNRHVSKSS
CNYNHHLDVVDNLTLMVEHTDIPEASPASTPQIIKHKALDLDDRWFKRSRLLDTDKRSKA
NTGSSNQDKASKMSSPETDEEIEKMKGFGEYSRSPTF

95/330

GGCAGACGCGCGGGCGGAGGGGCGAGAGTCAGCCGAGCCGAGTCCAGCCGGACGAGCGGACCAGCGCAGGGCAGC
CCAAGCAGCGCGCGCAGCGAACGCCCCGCCGCCACACCCCTCTGCGGTCCCCGCGGCGCCTGCCACCCTTCCCT
CCTTCCCCCGCGTCCCCGCGCTCGCCGGCCAGTCAGCTTGCCGGGTTTCGCTGCCCGCGGAAACCCCGAGGTCACCA
GCCCCGCGCCTCTGCTTCCCTGGGCGCGCGCGCCGCTCCACGCCCCTCCTTCTCCCCCTGGCCCCGGCGCCTGGCACC
GGGGACCGTTGCCTGACGCGAGGCCAGCTCTACTTTTTCGCCCCGCGTCTCCTCCGCGCTGCTCGCCTCTTCCAC
CAACTCCAACCTCCTTCTCCCTCCAGCTCCACTCGTAGTCCCCGACTCCGCGACCCCTCGGCCGCTGCCGTAG
CGCCGCTTCCCCGTCGGTCCCCAAGGTGGGAACGCGTCCGCCCCGCGCCGACCTGCGCACGGTTCGGCTTGCC
CGCGCTTCTCTGCACCTGGCAGTGCTCAGCGCCGCGCTGCTGGCTGCCGAGCTCAAGTCGAAAAGTTGCTCGG
AAGTGCAGACGCTCTTTACGTGTCCAAAGGCTTCAACAAGAACGATGCCCCCTCCACGAGATCAACGGTGATCAT
TTGAAGATCTGTCCCCAGGGTTCTACCTGCTGCTCTCAAGAGATGGAGGAGAAGTACAGCCTGCAAAGTAAAGA
TGATTTCAAAGTGTGGTCAGCGAACAGTGCAATCATTTGCAAGCTGTCTTTGCTTACGTTACAAGAAGTTTG
ATGAATTCCTTCAAAGAACTACTTGAAAATGCAGAGAAATCCCTGAATGATATGTTTGTAAGACATATGGCCAT
TTATACATGCAAATTTCTGAGCTACTTTAAAGATCTCTTCGTAGAGTTGAAACGTTACTACGTGGTGGGAAATGT
GAACCTGGAAGAAATGCTAAATGACTTCTGGGCTCGCCTCTGGAGCGGATGTTCCGCTGGTGAACCTCCAGT
ACCATTTCACAGATGAGTATCTGGAAATGTGTGAGCAAGTATACGGAGCAGCTGAAGCCCTTCGGAGATGTCCCT
CGCAAAATGAAGCTCCAGGTTACTCGTGCTTTTGTAGCAGCCCGTACTTTGCTCAAGGCTTAGCGGTTGCGGG
AGATGTGCTGAGCAAGGTCTCCGTGGTAAACCCACAGCCCAGTGTACCCATGCCCTGTTGAAGATGATCTACT
GCTCCCACTGCCGGGGTCTCGTGACTGTGAAGCCATGTTACAACCTACTGCTCAAACATCATGAGAGGCTGTTTG
GCCAACCAAGGGGATCTCGATTTTGAATGGAACAATTTTCATAGATGCTATGCTGATGGTGGCAGAGAGGCTAGA
GGGTCTCTTCAACATTGATTCGGTCATGGATCCCATCGATGTGAAGATTTCTGATGCTATTATGAACATGCAGG
ATAATAGTTGTTCAAGTGTCTCAGAAGGTTTTCCAGGGATGTGGACCCCCCAAGCCCTCCAGCTGGACGAATT
TCTCGTTCCATCTCTGAAAGTGCCCTCAGTGCTCGCTTCAGACCACATCACCCCGAGGAACGCCCAACCACAGC
AGCTGGCACTAGTTTGGACCGACTGGTTACTGATGTCAAGGAGAAACTGAAACAGGCCAAGAAATTCTGGTCTCT
CCCTTCCGAGCAACGTTTGAACGATGAGAGGATGGCTGCAGGAAACGGCAATGAGGATGACTGTTGGAATGGG
AAAGGCCAAAAGCAGGTACCTGTTTGAGTGCAGGAAATGGATTAGCCAACAGGGGCAACAACCCAGAGGTTCA
GGTTGACACCAGCAAACCAGACATACTGATCCTTCGTCAAATCATGGCTCTTCGATGATGACCAAGATGA
AGAATGCATACAATGGGAACGACGCTGGACTTCTTTGATATCAGTGATGAAAGTAGTGGAGAAGGAAGTGGAAGT
GGCTGTGAGTATCAGCAGTGCCCTTCAGAGTTTGATACAATGCCACTGACCATGCTGGGAAGAGTGCCAATGA
GAAAGCCGACAGTGCTGGTGTCCGTCTGGGGCACAGGCCTACCTCCTCACTGTCTTCTGCATCTTGTTCTCTGG
TTATGCAGAGAGAGTGGAGATTAATTTCTCAAACCTCTGAGAAAAGTGTTTCATCAAAAAGTTAAAAGGCACCAAGT
ATCACTTTTCTACCATCCTAGTGACTTTTGCTTTTTTAATGAATGGACAACAATGTACAGTTTTTACTATGTGGC
CACTGGTTTTAAGAAGTGCTGACTTTGTTTTCTCATTGACTTTTGGGAGGAAAAGGGACTGTGCATTGAGTTGGT
TCCTGCTCCCCCAAAACCATGTTAAACGTGGTCAACAGTGTAGGTACAGAACTATAGTTAGTTGTGCATTTGTGA
TTTTTACTACTCTATTATTTTGTATGTTTTTTTCTCATTTCGTTTGTGGGTTTTTTTTTCCAACCTGTGATCT
CGCTTGTGTTTTTACAAGCAAACAGGGTCCCTTCTTGGCACGTAACATGTACGTATTTCTGAAATATTAATA
GCTGTACAGAAGCAGGTTTTTATTTATCATGTTATCTTATTAAGAAAAAGCCCAAAAGC

FIGURE 98

CTCGCCCTCAAATGGGAACGCTGGCCTGGGACTAAAGCATAGACCACCAGGCTGAGTATCCT
GACCTGAGTCATCCCCAGGGATCAGGAGCCTCCAGCAGGGAACCTTCCATTATATTCTTCAA
GCAACTTACAGCTGCACCGACAGTTGCG**ATG**AAGTTCTAATCTCTTCCCTCCTCCTGTTGC
TGCCACTAATGCTGATGTCCATGGTCTCTAGCAGCCTGAATCCAGGGGTCGCCAGAGGCCAC
AGGGACCGAGGCCAGGCTTCTAGGAGATGGCTCCAGGAAGGCGGCCAAGAATGTGAGTGCAA
AGATTGGTTCCTGAGAGCCCCGAGAAGAAAATTCATGACAGTGTCTGGGCTGCCAAAGAAGC
AGTGCCCCTGTGATCATTTCAAGGGCAATGTGAAGAAAACAAGACACCAAAGGCACACAGA
AAGCCAAACAAGCATTCCAGAGCCTGCCAGCAATTTCTCAAACAATGTCAGCTAAGAAGCTT
TGCTCTGCCTTTG**TAG**GAGCTCTGAGCGCCCACTCTTCCAATTAAACATTCTCAGCCAAGAA
GACAGTGAGCACACCTACCAGACACTCTTCTTCTCCCACCTCACTCTCCCACTGTACCCACC
CCTAAATCATTCCAGTGCTCTCAAAAAGCATGTTTTTCAAGATCATTTTGTTTGTTGCTCTC
TCTAGTGTCTTCTTCTCTCGTCAGTCTTAGCCTGTGCCCTCCCCTTACCCAGGCTTAGGCTT
AATTACCTGAAAGATTCCAGGAACTGTAGCTTCCTAGCTAGTGTCAATTAACCTTAAATGC
AATCAGGAAAGTAGCAAACAGAAGTCAATAAATATTTTTAAATGTCAAAAAAAAAAAAAAAAAA

99/330

FIGURE 99

MKVLISSLLLLPLMLMSVSSSLNPGVARGHRDRGQASRRWLQEGGQECECKDWFLRAPRR
KFMTVSGLPKKQCPCDHFKGNVKKTRHQRRHRKPNKHSRACQQFLKQCQLRSFALPL

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

101/330

FIGURE 101

MAVLVLRRLTVVLGLLVLF LTCYADDKPDKPDDKPD DSGKDPKPDFPKFLSLLGTEIIENAVE
FILRSMRSTGFMEFDDNEGKHSSK

101/330

FIGURE 102

GGACGCCAGCGCCTGCAGAGGCTGAGCAGGGAAAAAGCCAGTGCCCCAGCGGAAGCACAGCT
CAGAGCTGGTCTGCC**ATG**GACATCCTGGTCCCCTCCTGCAGCTGCTGGTGCTGCTTCTTAC
CCTGCCCTGCACCTCATGGCTCTGCTGGGCTGCTGGCAGCCCCCTGTGCAAAAGCTACTTCC
CCTACCTGATGGCCGTGCTGACTCCCAAGAGCAACCGCAAGATGGAGAGCAAGAAACGGGAG
CTCTTCAGCCAGATAAAGGGGCTTACAGGAGCCTCCGGGAAAGTGGCCCTACTGGAGCTGGG
CTGCGGAACCGGAGCCAACTTTCAGTTCTACCCACCGGGCTGCAGGGTCACCTGCCTAGACC
CAAATCCCCACTTTGAGAAGTTCCTGACAAAGAGCATGGCTGAGAACAGGCACCTCCAATAT
GAGCGGTTTGTGGTGGCTCCTGGAGAGGACATGAGACAGCTGGCTGATGGCTCCATGGATGT
GGTGGTCTGCACTCTGGTGCTGTGCTCTGTGCAGAGCCCCAAGGAAGGTCTGCAGGAGGTCC
GGAGAGTACTGAGACCGGGAGGTGTGCTCTTTTTCTGGGAGCATGTGGCAGAACCATATGGA
AGCTGGGCCTTCATGTGGCAGCAAGTTTTTCGAGCCCACCTGGAAACACATTGGGGATGGCTG
CTGCCTCACCAGAGAGACCTGGAAGGATCTTGAGAACGCCCAGTTCTCCGAAATCCAAATGG
AACGACAGCCCCCTCCCTTGAAGTGGCTACCTGTTGGGCCCCACATCATGGGAAAGGCTGTC
AAACAATCTTTCCCAAGCTCCAAGGCACTCATTTGCTCCTTCCCCAGCCTCCAATTAGAACA
AGCCACCCACCAGCCTATCTATCTTCCACTGAGAGGGACCT**TAG**CAGAATGAGAGAAGACATT
CATGTACCACCTACTAGTCCCTCTCTCCCCAACCTCTGCCAGGGCAATCTCTAACTTCAATC
CCGCCTTCGACAGTGAAAAAGCTCTACTTCTACGCTGACCCAGGGAGGAAACACTAGGACCC
TGTTGTATCCTCAACTGCAAGTTTCTGGACTAGTCTCCCAACGTTTGCCTCCCAATGTTGTC
CCTTTCCTTCGTTCCCATGGTAAAGCTCCTCTCGCTTTCCTCCTGAGGCTACACCCATGCGT
CTCTAGGAAGTGGTCACAAAAGTCATGGTGCCTGCATCCCTGCCAAGCCCCCTGACCCTCT
CTCCCCACTACCACCTTCTTCTGAGCTGGGGGCACCAGGGAGAATCAGAGATGCTGGGGAT
GCCAGAGCAAGACTCAAAGAGGCAGAGGTTTTGTTCTCAAATATTTTTTAATAAATAGACGA
AACCACG

FIGURE 103

MDILVPLLQLLVLLLTLPPLHLMALLGCWQPLCKSYFPYLMAVLTPKSNRKMESKKRELFSQL
KGLTGASGKVALLELGCCTGANFQFYPPGCRVTCLDPNPHFEKFLTKSMAENRHLQYERFVV
APGEDMRQLADGSMDVVVCTLVLCVQSPRKVLQEVRRVLRPGGVLFWEHVAEPYGSWAFM
WQQVFEPTWKHIGDGCCLTRETWKDLENAQFSEIQMERQPPPLKWLPGPHIMGKAVKQSEF
SSKALICSFPSLQLEQATHQPIYLPLRGT

103/330

FIGURE 104

GTGGGATTTATTTGAGTGCAAGATCGTTTTCTCAGTGGTGGTGGAGTTGCCTCATCGCAGG
CAGATGTTGGGGCTTTGTCCGAACAGCTCCCCCTCTGCCAGCTTCTGTAGATAAGGGTTAAAA
ACTAATATTTATATGACAGAAGAAAAAG**ATG**TCATTCCGTAAAGTAAACATCATCATCTTGG
TCCTGGCTGTTGCTCTCTTCTTACTGGTTTTGCACCATAACTTCCTCAGCTTGAGCAGTTTG
TTAAGGAATGAGGTTACAGATTCAGGAATTGTAGGGCCTCAACCTATAGACTTTGTCCCAA
TGCTCTCCGACATGCAGTAGATGGGAGACAAGAGGAGATTCTGTGGTCATCGCTGCATCTG
AAGACAGGCTTGGGGGGGCCATTGCAGCTATAAACAGCATTGAGCACAACACTCGCTCCAAT
GTGATTTTCTACATTGTTACTCTCAACAATACAGCAGACCATCTCCGGTCTGGCTCAACAG
TGATTCCCTGAAAAGCATCAGATACAAAATTGTCAATTTTGACCCTAAACTTTTGAAGGAA
AAGTAAAGGAGGATCCTGACCAGGGGGAATCCATGAAACCTTTAACCTTTGCAAGGTTCTAC
TTGCCAATTCTGGTTCCCAGCGCAAAGAAGGCCATATACATGGATGATGATGTAATTGTGCA
AGGTGATATTCTTGCCCTTTACAATACAGCACTGAAGCCAGGACATGCAGCTGCATTTTCAG
AAGATTGTGATTGAGCCTCTACTAAAGTTGTCATCCGTGGAGCAGGAAACCAGTACAATTAC
ATTGGCTATCTTGACTATAAAAAGGAAAGAATTTCGTAAGCTTTCCATGAAAGCCAGCACTTG
CTCATTTAATCCTGGAGTTTTTGTGCAAACCTGACGGAATGGAAACGACAGAATATAACTA
ACCAACTGGAAAAATGGATGAACTCAATGTAGAAGAGGGACTGTATAGCAGAACCCTGGCT
GGTAGCATCACAAACCTCCTCTGCTTATCGTATTTTATCAACAGCACTCTACCATCGATCC
TATGTGGAATGTCCGCCACCTTGGTTCCAGTGCTGGAAAACGATATTCACCTCAGTTTGTAA
AGGCTGCCAAGTTACTCCATTGGAATGGACATTTGAAGCCATGGGGAAGGACTGCTTCATAT
ACTGATGTTTGGGAAAAATGGTATATTCCAGACCCAACAGGCAAATTC AACCTAATCCGAAG
ATATACCGAGATCTCAAACATAAAG**TGA**AACAGAATTTGAACTGTAAGCAAGCATTTCTCAG
GAAGTCCTGGAAGATAGCATGCATGGGAAGTAACAGTTGCTAGGCTTCAATGCCTATCGGTA
GCAAGCCATGGAAAAAGATGTGTCAGCTAGGTAAAGATGACAACTGCCCTGTCTGGCAGTC
AGCTTCCCAGACAGACTATAGACTATAAATATGTCTCCATCTGCCTTACCAAGTGTTCCTT
ACTACAATGCTGAATGACTGGAAAGAAGAACTGATATGGCTAGTTCAGCTAGCTGGTACAGA
TAATTCAAACTGCTGTTGGTTTTAATTTTGTAACTGTGGCCTGATCTGTAAATAAACTT
ACATTTTTC

105/330

FIGURE 105

MSFRKVNIIILVLAVALFLLVLHHNFLSLSSLLRNEVTD SGIVGPPIDFVPNALRHAVDGR
QEEIPVVIAASEDRLGGAIAAINSIQHNTRSNVIFYIVTLNNTADHLRSWLNSDSLKSIRYK
IVNFDPKLLEGKVKEPDQGESMKPLTFARFYLPILVPSAKKAIYMDDD VIVQG DILALYNT
ALKPGHAAAFSEDCDSASTKVVIRGAGNQYNYIGYLDYKKERIRKLSMKASTCSFNPGVFVA
NLTEWKRQNITNQLKWMKLNVEEGLYSRTL AGSITTPPLLIVFYQQHSTIDPMWNVRLGS
SAGKRYSPQFVKA AKLLHWNGHLKPWGRTASYTDVWEKWI PDPTGKFNLIRRYTEISNIK

105/330

FIGURE 106

TGGTTTTTGCCCCATAAATTCCTCAGCTTGAGCAGTTTGTTAAGGAATGAGGTTACAGATT
CAGGAATTNTAGGNCCTCAACCTNTAGANTTTGTCCCAAATGTTCTCCGACATGCAGTAGAT
GGGAGACAAGAGGAGATTCCCTGTGGTCATCGCTGCATNTGAAGACAGGCTTGGGGGGGCCAT
TGCAGCTATAAACAGCATTTCAGCACAACTCGNTCCAATGTGATTTTCTACATTGTTACTC
TCAACAATACAGCAGACCATNTCCGGTCCTGGNTCAACAGTGATTCCCTGAAAAGCATCAGA
TACAAAATTGTCAATTTTGACCCTAACTTTTGGAAGGAAAAGTAAAGGAGGATCCTGACCA
GGGGGAATCCATGAAACCTTTAACCTTTGCAAGGTTCTACTTGCCAATTCTGGTTCCCAGCG
CAAAGAAGGCCATATACATGGATGATGATGTAATTGTGCAAGGTGATATTCTTGCCCTTTAC
AATACAGCACTGAAGCCAGGACATGCAGCTGCATTTTCAGAAGATTGTGATTCAGCCTCTAC
TAAAGTTGTCATCCGTGGAGCAGGAAA

106/330

FIGURE 107

CGACGCTCTAGCGGTTACCGCTGCGGGCTGGCTGGGCGTAGTGGGGCTGCGCGGCTGCCACG
GAGCTAGAGGGCAAGTGTGCTCGGCCCAGCGTGCAGGGAACGCGGGCGGCCAGACAACGGGC
TGGGCTCCGGGGCCTGCGGGCGGGGCGCTGAGCTGGCAGGGCGGGTTCGGGGCGCGGGCTGCA
TCCGCATCTCCTCCATCGCCTGCAGTAAGGGCGGGCGGGCGAGCCTTTGAGGGGAACGACT
TGTCGGAGCCCTAACCAGGGGTGTCTCTGAGCCTGGTGGGATCCCCGGAGCGTCACATCACT
TTCCGATCACTTCAAAGTGGTTAAAACTAATATTTATATGACAGAAGAAAAAGATGTCATT
CCGTAAAGTAAACATCATCATCTTGGTCCTGGGCTGTTGCTCTCTTCTTACTGGTTTTGCAC
CATAACTTCCTCAGCTTGAGGCAGTTTGTAAAGGAATGAGGTTACAGATTCAGGAATTGTAG
GGCCTCAACCTATAGGACTTTGTCCCAAATGCTCTCCGACATGCAGTAGATGGGAGACAAGA
GGAGATTCCCTGTGGTCATCGCTGCATCTGAAGACAGGCTTGGGGGGGCCATTGCAGCTATAA
ACAGCATTACAGCACAACACTCGCTCCAATGTGATTTTCTACATTGTTACTCTCAACAATACA
GCAGACCATCTCCGGTCTTGGGCTCAACAGTGATTCCCTGAAAAGCATCAGATACAAAATTG
TCAATTTTGACCCTAACTTTTGAAGGAAAAGTAAAGGAGGATCCTGACCAGGGGGAATCC
ATGAAACCTTTAACCTTTGCAAGGTTCTACTTGCCAATTCTGGGTTCACAGCGCAAAGAAGG
CCATATACATGGATGATGATGTAATTGTGCAAGGTGATATTCTTGCCCTTTACAATACAGCA
CTGAAGCCAGGACATGCAGCTGCATTTTCAGAAGATTGTGATTACAGCCTCTACTAAAGTTGT
CATCCGTGGAGCAGGAAACCAGTACAATTACATTGGCTATCTTGACTATAAAAAGGAAAGAA
TTCGTAAGCTTTCCATGAAAGCCAGCACTTGCTCATTTAATCCTGGAGTTTTTGTGCAAAC
CTGACGGAATGGAACGACAGAATATAACTAACCAACTGGAAAAATGGATGAACTCAATGT
AGAAGAGGGGACTGTATAGCAGAACCCTGGCTGGTAGCATCACAAACACCTCCTCTGCTTATCG
TATTTTATCAACAGCACTCTACCATCGATCCTATGTGGAATGTCCGCCACCTTGGTTCCAGT
GCTGGAAAACGATATTCACCTCAGTTTGTAAAGGCTGCCAAGTTACTCCATTGGAATGGACA
TTTGAAGCCATGGGGAAGGACTGCTTCATATACTGATGTTTGGGGAAAAATGGTATATTCCA
GACCCAACAGGCAAATTCAACCTAATCCGAAGATATACCGAGATCTCAAACATAAAGTGAAA
CAGAATTTGAACTGTAAGCAAGCATTTCTCAGGAAGTCCTGGAAGATAGCATGCGTGGAAG
TAACAGTTGCTAGGCTTCAATGCCTATCGGTAGCAAGCCATGGAAAAAGATGTGTCAGCTAG
GTAAAGATGACAACTGCCCTGTCTGGCAGTCAGCTTCCCAGACAGACTATAGACTATAAAT
ATGTCTCCATCTGCCTTACCAAGTGTTTTCTTACTACAATGCTGAATGACTGGAAAGAAGAA
CTGATATGGCTAGTTCAGCTAGCTGGTACAGATAATTCAAACTGCTGTTGGTTTTAATTTT
GTAACCTGTGGCCTGATCTGTAAATAAACTTACATTTTCAATAGGTAAAAAAAAAAAAAA
AAAAAA

CGAGCTCTAGCGGTTACCGCTGCGGGCTGGCTGGGCGTAGTGGGGCTGCGCGGCTGCCACG

FIGURE 108

CTGCAGGTAGACATCTCCACTGCCCAGGAATCACTGAGCGTGCAGACAGCACAGCCTCCTCT
GAAGGCCGGCCATACCAGAGTCCTGCCTCGGCATGGGCCTCACCATTGAGGCAGCTCCACTG
TCTGTGCTGGTCTGAGGGTGCTGCCTGTCATGGGGGCAGCCATCTCCCAGGGGGCCCTCATC
GCCATCGTCTGCAACGGTCTCGTGGGCTTCTTGCTGCTGCTGCTCTGGGTCATCCTCTGCTG
GGCCTGCCATTCTCGTCTGCCGACGTTGACTCTCTCTCTGAATCCAGTCCCAACTCCAGCCC
TGGCCCCTGTCCTGAGAAGGCCCCACCACCCCAGAAGCCCAGCCATGAAGGCAGCTACCTGC
TGCAGCCCTGAAGGCCCTGGCCTAGCCTGGAGCCCAGGACCTAAGTCACCTCACCTAGAG
CCTGGAATTAGGATCCCAGAGTTCAGCCAGCCTGGGGTCCAGAACTCAAGAGTCCGCCTGCT
TGGAGCTGGACCCAGCGGCCCAGAGTCTAGCCAGCTTGGCTCCAATAGGAGCTCAGTGGCCC
TAAGGAGATGGGCCTGGGGTGGGGGCTTATGAGTTGGTGCTAGAGCCAGGGCCATCTGGACT
ATGCTCCATCCCAAGGGCCAAGGGTCAGGGGCCGGGTCCACTCTTTCCCTAGGCTGAGCACC
TCTAGGCCCTCTAGGTTGGGGAAGCAAACCTGGAACCCATGGCAATAATAGGAGGGTGTCCAG
GCTGGGGCCCCTCCCCTGGTCCTCCCAGTGTTTGCTGGATAATAAATGGAACCTATGGCTCTAA
AAAAAAAAAAAAAAAAAAAA

108/330

109/330

FIGURE 109

MGAAISQGALIAIVCNGLVGFLLLLLWVILCWACHSRLPTLTLSLNPVPTPALAPVLRPHH
PRSPAMKAATCCSPEGPWPSLEPRT

109/330

FIGURE 110

GTTTGAATTCCCTTCAACTATACCCACAGTCCAAAAGCAGACTCACTGTGTCCCAGGCTACCA
 GTTCCTCCAAGCAAGTCATTTCCCTTATTTAACCGATGTGTCCCTCAAACACCTGAGTGCTA
 CTCCTTATTTGCATCTGTTTTGATAAATGATGTTGACACCCTCCACCGAATTCTAAGTGGA
 TC**ATG**TCGGGAAGAGATACAATCCTTGGCCTGTGTATCCTCGCATTAGCCTTGTCTTTGGCC
 ATGATGTTTACCTTCAGATTCATCACCACCCTTCTGGTTCACATTTTCATTTCAATTGGTTAT
 TTTGGGATTGTTGTTTGTCTGCGGTGTTTTATGGTGGCTGTATTATGACTATACCAACGACC
 TCAGCATAGAATTGGACACAGAAAGGGAAAATATGAAGTGCGTGCTGGGGTTTGCTATCGTA
 TCCACAGGCATCACGGCAGTGCTGCTCGTCTTGATTTTTGTTCTCAGAAAGAGAATAAAATT
 GACAGTTGAGCTTTTCCAAATCACAAATAAAGCCATCAGCAGTGCTCCCTTCCTGCTGTTCC
 AGCCACTGTGGACATTTGCCATCCTCATTTTCTTCTGGGTCTCTGGGTGGCTGTGCTGCTG
 AGCCTGGGAAGTGCAGGAGCTGCCAGGTTATGGAAGGCGGCCAAGTGGAATATAAGCCCCT
 TTCGGGCATTTCGGTACATGTGGTCGTACCATTTAATTGGCCTCATCTGGACTAGTGAATTCA
 TCCTTGCGTGCCAGCAAATGACTATAGCTGGGGCAGTGGTTACTTGTTATTTCAACAGAAGT
 AAAAATGATCCTCCTGATCATCCCATCCTTTTCGTCTCTCTCCATTCTCTTCTTCTACCATCA
 AGGAACCGTTGTGAAAGGGTCATTTTTAATCTCTGTGGTGAGGATTCCGAGAATCATTGTC
 TGTACATGCAAAACGCACTGAAAGAACAGCAGCATGGTGCATTGTCCAGGTACCTGTTCCGA
 TGCTGCTACTGCTGTTTCTGGTGTCTTGACAAATACCTGCTCCATCTCAACCAGAATGCATA
 TACTACAACTGCTATTAATGGGACAGATTTCTGTACATCAGCAAAAGATGCATTCAAAATCT
 TGTTCAAGAACTCAAGTCACTTTACATCTATTAAGTCTTTGGAGACTTCATAATTTTCTA
 GGAAAGGTGTTAGTGGTGTGTTTCACTGTTTTTGGAGGACTCATGGCTTTTAACTACAATCG
 GGCATTCCAGGTGTGGGCAGTCCCTCTGTTATTGGTAGCTTTTTTTGCCTACTTAGTAGCCC
 ATAGTTTTTTATCTGTGTTTGAAACTGTGCTGGATGCACTTTTCTGTGTTTTGCTGTGAT
 CTGGAAACAAATGATGGATCGTCAGAAAAGCCCTACTTTATGGATCAAGAATTTCTGAGTTT
 CGTAAAAAGGAGCAACAAATTAACAATGCAAGGGCACAGCAGGACAAGCACTCATTAAGGA
 ATGAGGAGGGAACAGAACTCCAGGCCATTGTGAGAT**TAG**ATACCCATTTAGGTATCTGTACCT
 GGAAACATTTTCCTTCTAAGAGCCATTTACAGAATAGAAGATGAGACCACTAGAGAAAAGTT
 AGTGAATTTTTTTTTTAAAAGACCTAATAAACCCCTATTCTTCCTCAAAA

FIGURE 111

MSGRDTILGLCILALALSLAMMFTFRFITLLVHIFISLVILGLLFVCGVLWWLYDYTNDL
SIELDTERENMKCVLGFAIVSTGITAVLLVLI FVLRKRIKLTVELFQITNKAISSAPFLLFQ
PLWTFAILIFFWVLWVAVLLSLGTAGAAQVMEGGQVEYKPLSGIRYMWSYHLIGLIWTSEFI
LACQQMTIAGAVVTCYFNRSKNDPPDHPILSSLSILFFYHQGTVVKGSFLISVVRIPRIIVM
YMQNALKEQQHGALSRYLFRCCYCCFWCLDKYLLHLNQNAYTTTAINGTDFCTSAKDAFKIL
SKNSSHFTSINCFGDFIIFLGKVLVVCFTVFGGLMAFNYNRAFQVWAVPLLLVAFFAYLVAH
SFLSVFETVLDALFLCFAVDLETNDGSSEKPYFMDQEFLSFVKRSNKLNNARAQQDKHSLRN
EEGTELQAIVR

111/330

[illegible]

FIGURE 113

MRTVVLTMKASVIEMFLVLLVTGVHSNKETAKKIKRPKFTVPQINCDVKAGKIIDPEFIVKC
PAGCQDPKYHVGTDVYASYSSVCGAAVHSGVLDNSGGKILVRKVAGQSGYKGSYSNGVQSL
SLPRWRESFIVLESKPKKGVTPSALTYSKSPAAQAGETTKAYQRPIPGTTAQPVTLMQ
LLAVTVAVATPTTLPRPSPSAASTTSIPRPQSVGHRSEQEMDLWSTATYTSSQNRPRADPGIQ
RQDPGAAAFQKPVGADVSLGLVPKEELSTQSLEPVS LGDPNCKIDLSFLIDGSTSIGKRRFR
IQKQLLADVAQALDIGPAGPLMGVVQYGDNPATHFNLKTHTNRDLKTAIEKITQRGGLSNV
GRAISFVTKNFFSKANGNRSGAPNVVVVMVDGWPTDKVEEASRLARESGINIFFITIEGAAE
NEKQYVVEPNFANKAVCRTNGFYSLHVQSWFGLHKT LQPLVKRVCDTDRLACSKTCLNSADI
GFVIDGSSSVGTGNFRTVLQFVTNLTKEFEISDTDTRIGAVQYTYEQRLEFGFDKYSSKPD I
LNAIKRVGYWSGGTSTGAAINFALEQLFKKSKPNKRKLMILITDGRSYDDVRI PAMAAHLKG
VITYAIGVAWAAQEELEVIATHPARDHSFFVDEFDNLHQYVPRIIQNICTEFNSQPRN

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 114

CAGGATGAACTGGTTGCAGTGGCTGCTGCTGCTGCGGGGGCGCTGAGAGGACACGAGCTCTA
TGCCTTTCCGGCTGCTCATCCCGCTCGGCCTCCTGTGCGCGCTGCTGCCTCAGCACCATGGT
GCGCCAGGTCCCGACGGCTCCGCGCCAGATCCCGCCCCACTACAGTTTTTCTCTGACTCTAAT
TGATGCACTGGACACCTTGCTGATTTTGGGGAATGTCTCAGAATCCAAAGAGTGGTTGAAG
TGCTCCAGGACAGCGTGGACTTTGATATTGATGTGAACGCCTCTGTGTTTGAAACAAACATT
CGAGTGGTAGGAGGACTCCTGTCTGCTCATCTGCTCTCCAAGAAGGCTGGGGTGGAAGTAGA
GGCTGGATGGCCCTGTTCCGGGCCTCTCCTGAGAATGGCTGAGGAGGCGGCCCCGAAAACCTCC
TCCCAGCCTTTCAGACCCCCACTGGCATGCCATATGGAACAGTGAACCTTACTTCATGGCGTG
AACCCAGGAGAGACCCCTGTCACCTGTACGGCAGGGATTGGGACCTTCATTGTTGAATTTGC
CACCTGAGCAGCCTCACTGGTGACCCGGTGTTCGAAGATGTGGCCAGAGTGGCTTTGATGC
GCCTCTGGGAGAGCCGGTCAGATATCGGGCTGGTTCGCAACCACATTGATGTGCTCACTGGC
AAGTGGGTGGCCCAGGACGCAGGCATCGGGGCTGGCGTGGACTCCTACTTTGAGTACTTGGT
GAAAGGAGCCATCCTGCTTCAGGATAAGAAGCTCATGGCCATGTTCCCTAGAGTATAACAAAG
CCATCCGGAACACACCCGCTTCGATGACTGGTACCTGTGGGTTTCAAGTGTACAAGGGGACT
GTGTCCATGCCAGTCTTCCAGTCCCTTGAGGGCCTACTGGCCTGGTCTTCAGAGCCTCATTGG
AGACATTGACAATGCCATGAGGACCTTCCTCAACTACTACACTGTATGGAAGCAGTTTGGGG
GGCTCCCGGAATTCTACAACATTCTCAGGGATACACAGTGGAGAAGCGAGAGGGCTACCCA
CTTCGGCCAGAACTTATTGAAAGCGCAATGTACCTCTACCGTGCCACGGGGGATCCCACCT
CCTAGAACTCGGAAGAGATGCTGTGGAATCCATTGAAAAAATCAGCAAGGTGGAGTGCGGAT
TTGCAACAATCAAAGATCTGCGAGACCACAAGCTGGACAACCGCATGGAGTCGTTCTTCCTG
GCCGAGACTGTGAAATACCTCTACCTCCTGTTTGACCCAACCAACTTCATCCACAACAATGG
GTCCACCTTCGACGCGGTGATCACCCCTATGGGGAGTGCATCCTGGGGGCTGGGGGGTACA
TCTTCAACACAGAAGCTCACCCCATCGACCTTGCCGCCCTGCACTGCTGCCAGAGGCTGAAG
GAAGAGCAGTGGGAGGTGGAGGACTTGATGAGGGAATTCTACTCTCTCAAACGGAGCAGGTC
GAAATTCAGAAAAACACTGTTAGTTCGGGGCCATGGGAACCTCCAGCAAGGCCAGGAACAC
TCTTCTCACCAGAAAACCATGACCAGGCAAGGGAGAGGAAGCCTGCCAAACAGAAGGTCCCA
CTTCTCAGCTGCCCCAGTCAGCCCTTACCTCCAAGTTGGCATTACTGGGACAGGTTTTCT
AGACTCCTCATAACCACTGGATAATTTTTTTATTTTTTATTTTTTTGAGGCTAAACTATAATA
AATTGCTTTTGCTATCATAAAA

FIGURE 115

MPFRLLIPLGLLCALLPQHHGAPGPDGSAPDPAHYSFSLTLIDALDTLLILGNVSEFQRVVE
VLQDSVDFDIDVNASVFETNIRVVGLLSAHLLSKKAGVEVEAGWPCSGPLLMAEEAARKL
LPAFQTPTGMPYGTVNLLHGVNPGETPVTCTAGIGTFIVEFATLSSLTGDPVFFEDVARVALM
RLWESRSDIGLVGNHIDVLTGKWVAQDAGIGAGVDSYFEYLVKGAILLQDKKLMAMFLEYNK
AIRNYTRFDDWYLVWQMYKGTVSMVPVQSLEAYWPGLQSLIGDIDNAMRTFLNYYTVWKQFG
GLPEFYNI PQGYTVEKREGYPLRPELIESAMYLYRATGDPTLLELGRDAVESIEKISKVECG
FATIKDLRDHKLDNRMESFFLAETVKYLYLLFDPTNFIHNNGSTFDAVITPYGECILGAGGY
IFNTEAHPIDLAALHCCQRLKEEQWEVEDLMREFYSLKRSRSKFQKNTVSSGPWEPPARPGT
LFSPENHDQARERKPAKQKVPLLSCPSQPFTSKLALLGQVFLDSS

115/330

FIGURE 116

AAAGTTACATTTTCTCTGGAACCTCTCCTAGGCCACTCCCTGCTGATGCAACATCTGGGTTTG
GGCAGAAAGGAGGGTGCTTCGGAGCCCGCCCTTTCTGAGCTTCCTGGGCCGGCTCTAGAACA
ATTGAGGCTTCGCTGCGACTCAGACCTCAGCTCCAACATATGCATTCTGAAGAAAGATGGCT
GAGATGGACAGAATGCTTTATTTTGGAAAGAAACAATGTTCTAGGTCAAACCTGAGTCTACCA
AAATGCAGACTTTCACAATGGTTCTAGAAGAAATCTGGACAAGTCTTTTCATGTGGTTTTTCT
ACGCATTGATTCCATGTTTGCTCACAGATGAAGTGGCCATTCTGCCTGCCCCCTCAGAACCTC
TCTGTACTCTCAACCAACATGAAGCATCTCTTGATGTGGAGCCCAGTGATCGCGCCTGGAGA
AACAGTGTACTATTCTGTGCAATACCAGGGGGAGTACGAGAGCCTGTACACGAGCCACATCT
GGATCCCCAGCAGCTGGTGCTCACTCACTGAAGGTCCTGAGTGTGATGTCACTGATGACATC
ACGGCCACTGTGCCATACAACCTTCGTGTCAGGGCCACATTGGGCTCACAGACCTCAGCCTG
GAGCATCCTGAAGCATCCCTTTAATAGAACTCAACCATCCTTACCCGACCTGGGATGGAGA
TCACCAAAGATGGCTTCCACCTGGTTATTGAGCTGGAGGACCTGGGGCCCCAGTTTGAGTTC
CTTGTTGGCCTACTGGAGGAGGGAGCCTGGTGCCGAGGAACATGTCAAAATGGTGAGGAGTGG
GGGTATTCCAGTGCACCTAGAAACCATGGAGCCAGGGGCTGCATACTGTGTGAAGGCCCAGA
CATTCGTGAAGGCCATTGGGAGGTACAGCGCCTTCAGCCAGACAGAAATGTGTGGAGGTGCAA
GGAGAGGCCATTCCCCCTGGTACTGGCCCTGTTTGCCTTTGTTGGCTTCATGCTGATCCTTGT
GGTCGTGCCACTGTTTCGTCTGGAAAATGGGCCGGCTGCTCCAGTACTCCTGTTGCCCCGTGG
TGGTCCCTCCCAGACACCTTGAAAATAACCAATTCACCCAGAAAGTTAATCAGCTGCAGAAGG
GAGGAGGTGGATGCCTGTGCCACGGCTGTGATGTCTCCTGAGGAACTCCTCAGGGCCTGGAT
CTCA**TAG**GTTTTCGGAAGGGCCCAGGTGAAGCCGAGAACCTGGTCTGCATGACATGGAAACC
ATGAGGGGGACAAGTTGTGTTTTCTGTTTTCCGCCACGGACAAGGGATGAGAGAAGTAGGAAGA
GCCTGTTGTCTACAAGTCTAGAAGCAACCATCAGAGGCAGGGTGGTTTGTCTAACAGAACAC
TGACTGAGGCTTAGGGGATGTGACCTCTAGACTGGGGGCTGCCACTTGCTGGCTGAGCAACC
CTGGGAAAAGTGACTTCATCCCTTCGGTCCCTAAGTTTTCTCATCTGTAATGGGGGAATTACC
TACACACCTGCTAAACACACACACACAGAGTCTCTCTCTATATATACACACGTACACATAAA
TACACCCAGCACTTGCAAGGCTAGAGGGAACTGGTGACACTCTACAGTCTGACTGATTCAG
TGTTTCTGGAGAGCAGGACATAAATGTATGATGAGAATGATCAAGGACTCTACACACTGGGT
GGCTTGGAGAGCCCACTTTCACAGAATAATCCTTGAGAGAAAAGGAATCATGGGAGCAATGG
TGTTGAGTTCACCTCAAGCCCAATGCCGGTGACAGAGGGGAATGGCTTAGCGAGCTCTACAGT
AGGTGACCTGGAGGAAGGTCACAGCCACACTGAAAATGGGATGTGCATGAACACGGAGGATC
CATGAACTACTGTAAAGTGTGACAGTGTGTGCACACTGCAGACAGCAGGTGAAATGTATGT
GTGCAATGCGACGAGAATGCAGAAGTCAGTAACATGTGCATGTTTGTGTGCTCCTTTTTTCT
TGTTGGTAAAGTACAGAATTACAGCAAATAAAAAGGGCCACCCTGGCCAAAAGCGGTAAAAAA
AAAAA

FIGURE 117

MQTFTMVLEEIWTSLEFMWFFYALIPCLLTDEVAILPAPQNLSVLSTNMKHLMLWSPVIAPGE
TVYYSVEYQGEYESLYTSHIWIPSSWCSLTEGPECDVTDDITATVPYNLRVRATLGSQTS
SILKHPFNRNSTILTRPGMEITKDGFLVIELEDLGPQFEFLVAYWRREPAAEEHVKMVRSG
GIPVHLETMEPGAAYCVKAQTFVKAIGRYSAFSQTECVEVQGEAIPLVLALFAFVGFMILILV
VVPLFVWKMGRLLQYSCCPVVLPDTLKITNSPQKLISCRREEVDACATAVMSPEELLRAWIS

Important features:

Signal peptide:

amino acids 1-29

Transmembrane domain:

amino acids 230-255

N-glycosylation sites.

amino acids 40-43 and 134-137

Tissue factor proteins homology.

amino acids 92-119

Integrins alpha chain protein homology.

amino acids 232-262

117/330

FIGURE 118

TCCTGCTGATGCACATCTGGGTTTGGCAAAGGAGGTTGCTTCGAGCCGCCCTTTCTAGCTT
CCTGGCCGGCTCTAGAACAAATTCAGGCTTCGCTGCGACTAGACCTCAGCTCCAACATATGCA
TTCTGAAGAAAGATGGCTGAGATGACAGAATGCTTTATTTTGGAAAGAAACAATGTTCTAGG
TCAAACCTGAGTCTACCAAATGCAGACTTTCACAATGGTTCTAGAAGAAATCTGGACAAGTCT
TTTCATGTGGTTTTTCTACGCATTGATTCCATGTTTGCTCACAGATGAAGTGGCCATTCTGC
CTGCCCCCTCAGAACCTCTCTGTACTCTCAACCAACATGAAGCATCTCTTGATGTGGAGCCCA
GTGATCGCGCCTGGAGAAACAGTGTACTATTCTGTGCGAATACCAGGGGGAGTACGAGAGCCT
GTACACGAGCCACATCTGGATCCCCAGCAGCTGGTGCTCACTCACTGAAGGTCCTGAGTGTG
ATGTCACTGATGACATCACGGCCACTGTGCCATACAACCTTTGTGTCAGGGCCACATTGGGC
TCACAGACCTCAGCCTGGAGCATCCTGAAGCATCCCTTTAATAGAACTCAACCATCCTTAC
CCGACCTGGGATGGAGATCACCAAAGATGGCTTNCACCTGGTTATTGAGCTGGAGGACCTGG
GGCCCCAGTTTGAGTTCCTTGTGGCCTANTGGAGGAGGGGCGAACCCCTTGCGGCGCAAGGG
GTTNGCGAACCCCTTGCGGCCGCTGGGGTATCTCTCGAGAAAAGAGAGGCCCAATATGACCCAC
ATACTCAATATGGACGAANTGCTATTGTCCACCTGTTTGAGTGGCGCTGGGTTGAT

CGGACGCGTGGGGCCGCCACCTCCGGAACAAGCC**ATG**GTGGCGGCGACGGTGGCAGCGGGCGTG
GCTGCTCCTGTGGGCTGCGGCCTGCGCGCAGCAGGAGCAGGACTTCTACGACTTCAAGGCGG
TCAACATCCGGGGCAAACCTGGTGTGCTGGAGAAGTACCGCGGATCGGTGTCCCTGGTGGTG
AATGTGGCCAGCGAGTGCGGCTTCACAGACCAGCACTACCGAGCCCTGCAGCAGCTGCAGCG
AGACCTGGGCCCCCACCACCTTTAACGTGCTCGCCTTCCCCTGCAACCAGTTTGGCCAACAGG
AGCCTGACAGCAACAAGGAGATTGAGAGCTTTGCCCCGCCGACCTACAGTGTCTCATTCCCC
ATGTTTAGCAAGATTGCAGTCACCGGTACTGGTGCCCATCCTGCCTTCAAGTACCTGGCCCCA
GACTTCTGGGAAGGAGCCCCACCTGGAACCTTCTGGAAGTACCTAGTAGCCCCAGATGGAAAGG
TGGTAGGGGCTTGGGACCCAACTGTGTCAGTGGAGGAGGTGAGACCCAGATCACAGCGCTC
GTGAGGAAGCTCATCCTACTGAAGCGAGAAGACTTAT**TAA**CCACCGCGTCTCCTCCTCCACCA
CCTCATCCCGCCACCTGTGTGGGGCTGACCAATGCAAACCTCAAATGGTGCTTCAAAGGGAG
AGACCCACTGACTCTCCTTCCCTTTACTCTTATGCCATTGGTCCCATCATTCTTGTGGGGGAA
AAATTCTAGTATTTTGATTATTTGAATCTTACAGCAACAAATAGGAACTCCTGGCCAATGAG
AGCTCTTGACCAGTGAATCACCAGCCGATACGAACGTCTTGCCAACAAAAATGTGTGGCAAA
TAGAAGTATATCAAGCAATAATCTCCCACCCAAGGCTTCTGTAAACTGGGACCAATGATTAC
CTCATAGGGCTGTTGTGAGGATTAGGATGAAATACCTGTGAAAGTGCCTAGGCAGTGCCAGC
CAAATAGGAGGCATTCAATGAACATTTTTTGCATATAAACCAAAAAATAACTTGTATCAAT
AAAAACTTGCATCCAACATGAATTTCCAGCCGATGATAATCCAGGCCAAAGGTTTAGTTGTT
GTTATTTCCCTCTGTATTATTTTCTTCATTACAAAAGAAATGCAAGTTCATTGTAACAATCCA
AACCAATACCTCAGGATATAAAAAATAAAAAATGAAAGTATCCTCCTCAAAAA

FIGURE 121

CGGACGCGTG GGGCGGGCCGGGACGCAGGGCAAAGCGAGCC**ATG**GCTGTCTACGTCGGGATGC
 TGC GCCTGGGGAGGCTGTGCGCCGGGAGCTCGGGGGTGCTGGGGGGCCGGGGCCGCCCTCTCT
 CGGAGTTGGCAGGAAGCCAGGTTGCAGGGTGTCCGCTTCCTCAGTTCCAGAGAGGTGGATCG
 CATGGTCTCCACGCCCATCGGAGGCCTCAGCTACGTTCAAGGGTGCACCAAAAAGCATCTTA
 ACAGCAAGACTGTGGGCCAGTGCCTGGAGACCACAGCACAGAGGGTCCCAGAACGAGAGGCC
 TTGGTCGTCTCCATGAAGACGTCAGGTTGACCTTTGCCCAACTCAAGGAGGAGGTGGACAA
 AGCTGCTTCTGGCCTCCTGAGCATTGGCCTCTGCAAAGGTGACCGGCTGGGCATGTGGGGAC
 CTAACCTCCTATGCATGGGTGCTCATGCAGTTGGCCACCGCCAGGCGGGCATCATTTCTGGTG
 TCTGTGAACCCAGCCTACCAGGCTATGGAAGTGGAGTATGTCCTCAAGAAGGTGGGCTGCAA
 GGCCCTTGTGTTCCCAAGCAATTCAAGACCCAGCAATACTACAACGTCCTGAAGCAGATCT
 GTCCAGAAGTGGAGAATGCCAGCCAGGGGGCCTTGAAGAGTCAGAGGCTCCCAGATCTGACC
 ACAGTCATCTCGGTGGATGCCCTTTGCCGGGGACCTGCTCCTGGATGAAGTGGTGGCGGC
 TGGCAGCACACGGCAGCATCTGGACCAGCTCCAATACAACCAGCAGTTTCTGTCTGCCATG
 ACCCATCAACATCCAGTTCACCTCGGGGACAACAGGCAGCCCCAAGGGGGCCACCCTCTCC
 CACTACAACATTGTCAACAACTCCAACATTTTAGGAGAGCGCCTGAAACTGCATGAGAAGAC
 ACCAGAGCAGTTGCGGATGATCCTGCCCAACCCCTGTACCATTGCCTGGGTTCCGTGGCAG
 GCACAATGATGTGTCTGATGTACGGTGCCACCCTCATCCTGGCCTCTCCCATCTTCAATGGC
 AAGAAGGCACTGGAGGCCATCAGCAGAGAGAGAGGCACCTTCCTGTATGGTACCCCCACGAT
 GTTCGTGGACATTCTGAACCAGCCAGACTTCTCCAGTTATGACATCTCGACCATGTGTGGAG
 GTGTCAATTGCTGGGTCCCCTGACCTCCAGAGTTGATCCGAGCCATCATCAACAAGATAAAT
 ATGAAGGACCTGGTGGTTGCTTATGGAACCACAGAGAACAGTCCCGTGACATTTCGCGCACTT
 CCCTGAGGACACTGTGGAGCAGAAAGGCAGAAAGCGTGGGCAGAAATTATGCCTCACACGGAGG
 CCCGGATCATGAACATGGAGGCAGGGACGCTGGCAAAGCTGAACACGCCCGGGGAGCTGTGC
 ATCCGAGGGTACTGCGTCATGCTGGGCTACTGGGGTGAGCCTCAGAAGACAGAGGAAGCAGT
 GGATCAGGACAAGTGGTATTGGACAGGAGATGTCGCCACAATGAATGAGCAGGGCTTCTGCA
 AGATCGTGGGCCGCTCTAAGGATATGATCATCCGGGGTGGTGAGAACATCTACCCCGCAGAG
 CTCGAGGACTTCTTTCACACACACCCGAAGGTGCAGGAAGTGCAGGTGGTGGGAGTGAAGGA
 CGATCGGATGGGGGAAGAGATTTGTGCCTGCATTGCGCTGAAGGACGGGGAGGAGACCACGG
 TGGAGGAGATAAAAGCTTTCTGCAAAGGGAAGATCTCTCACTTCAAGATTCCGAAGTACATC
 GTGTTTGTCAAACTACCCCTCACCATTTCAGGAAAGATCCAGAAATTCAACTTCGAGA
 GCAGATGGAACGACATCTAAATCT**GTGA**ATAAAGCAGCAGGCCTGTCCTGGCCGGTGGCTT
 GACTCTCTCCTGTCAGAAATGCAACCTGGCTTTATGCACCTAGATGTCCCCAGCACCCAGTTC
 TGAGCCAGGCACATCAAAATGTCAAGGAATTGACTGAACGAACTAAGAGCTCCTGGATGGGTC
 CGGGAACCTCGCCTGGGCACAAGGTGCCAAAAGGCAGGCAGCCTGCCAGGCCCTCCCTCCTG
 TCCATCCCCACATTCCCCTGTCTGTCCTTGTGATTTGGCATAAAGAGCTTCTGTTTTCTTT
 GAAAAAAAAAAAAAAAAA

FIGURE 122

MAVYVGMLRLGRLCAGSSGVLGARAALSRWQEARLQGVRFSSREVD RMVSTPIGGLSYVQ
GCTKKHLNSKTVGQCLETTAQRVPEREALVVLHEDVRLTFAQLKEEVDKAASGLLSIGLCKG
DRLGMWGPNSYAWVLMQLATAQAGIILVSVNPAYQAMELEYVLKKVGCKALVFPKQFKTQQY
YNVLKQICPEVENAQPGALKSQRLPDLTTVISVDAPLP GTLLLDEVVAAGSTRQHLDQLQYN
QQFLSCHDPINIQFTSGTTGSPKGATLSHYNIVNNSNILGERLKLHEKTPEQLRMILPNPLY
HCLGSVAGTMMCLMYGATLILASPIFNGKKALEAISRERGTFLYGTPTMFVDILNQPDFSSY
DISTMCGGVIAGSPAPPELIRAIINKINMKDLVVAYGTTENS PVTFAHFPEDTVEQKAESVG
RIMPHTEARIMNMEAGTLAKLNTPGELCIRGYCVMLGYWGEPQKTEEAVDQDKWYWTGDVAT
MNEQGFCIKIVGRSKDMIIRGGENIYPAELEDFHHTHPKVQEVQVVGVKDDRMGEEICACIRL
KDGEETTVEEIKAFCKGKISHFKIPKYIVFVTNYPLTISGKIQKFKLREQMERHLNL

Signal Peptide:

amino acids 1-22

Transmembrane Domains:

amino acids 140-161, 213-229, 312-334

Putative AMP-binding Domain Signature:

amino acids 260-271

N-myristoylation Sites:

amino acids 19-24, 22-27, 120-125, 203-208, 268-273, 272-277,
314-319, 318-323, 379-384, 380-385, 409-413

N-glycosylation Site:

amino acids 282-285

FIGURE 123

CAACTCCAACATTTTAGGAGAGCGCCTGAAACTGCATGAGAAGACACCAGAGCAGTTGCGGA
TGATCCTGCCCCAACCCCCTGTACCATTGCCTGGGTTCCGTGGCAGGCACAATGATGTGTCTG
ATGTACGGTGCCACCCTCATCCTGGCCTCTCCCATCTTCAATGGCAAGAAGGCACTGGAGGC
CATCAGCAGAGAGAGAGGCACCTTCCTGTATGGTACCCCCACGATGTTTCGTGGACATTCTGA
ACCAGCCAGACTTCTCCAGTTATGACATCTCGACCATGTGTGGAGGTGTCATTGCTGGGTCC
CCTGCACCTCCAGAGTTGATCCGAGCCATCATCAACAAGATAAATATGAAGGACCTGGTGGT
TGCTTATGGAACCAACAGAGAACAGTCCCGTGACATTCGCGCACTTCCCTGAGGACACTGTGG
AGCAGAAGGCAGAAAGCGTGGGCAGAATTATGCCTCACACGGAGGCGCGGATCATGAACATG
GAGGCAGGGACGCTGGCAAAGCTGAACACGCCCCGGGGAGCTGTGCATCCGAGGGTACTGCGT
CATGCTGGGCTACTGGGGTGAGCCTCAGAAGACAGAGGAAGCAGTGGATCAGGACAAGTGGT
ATTGGACAGGAGATGTCGCCAC

123/330

FIGURE 124

GAGCAGGACGGAGCCATGGACCCCGCCAGGAAAGCAGGTGCCCAGGCCATGATCTGGACTGC
AGGCTGGCTGCTGCTGCTGCTGCTTCGCGGAGGAGCGCAGGCCCTGGAGTGCTACAGCTGCG
TGCAGAAAGCAGATGACGGATGCTCCCCGAACAAGATGAAGACAGTGAAGTGCGCGCCGGGC
GTGGACGTCTGCACCGAGGCCGTGGGGGCGGTGGAGACCATCCACGGACAATTCTCGCTGGC
AGTGCGGGGTTGCGGTTCTGGGACTCCCCGGCAAGAATGACCGCGGCCTGGATCTTCACGGGC
TTCTGGCGTTTCATCCAGCTGCAGCAATGCGCTCAGGATCGCTGCAACGCCAAGCTCAACCTC
ACCTCGCGGGCGCTCGACCCGGCAGGTAATGAGAGTGCATACCCGCCAACGGCGTGGAGTG
CTACAGCTGTGTGGGCCTGAGCCGGGAGGCGTGCCAGGGTACATCGCCGCCGGTCTGTGAGCT
GCTACAACGCCAGCGATCATGTCTACAAGGGCTGCTTCGACGGCAACGTCACCTTGACGGCA
GCTAATGTGACTGTGTCCTTGCCTGTCCGGGGCTGTGTCCAGGATGAATTCTGCACTCGGGA
TGGAGTAACAGGCCCAGGGTTCACGCTCAGTGGCTCCTGTTGCCAGGGGTCCCGCTGTAAC
CTGACCTCCGCAACAAGACCTACTTCTCCCCTCGAATCCCACCCCTTGTCCGGCTGCCCCCT
CCAGAGCCCACGACTGTGGCCTCAACCACATCTGTCACTTCTACCTCGGCCCCAGTGAG
ACCCACATCCACCACCAAAACCCATGCCAGCGCCAACCAGTCAGACTCCGAGACAGGGAGTAG
AACACGAGGCCTCCCGGGATGAGGAGCCCAGGTTGACTGGAGGCGCCGCTGGCCACCAGGAC
CGCAGCAATTCAGGGCAGTATCCTGCAAAAGGGGGCCCCAGCAGCCCCATAATAAAGGCTG
TGTGGCTCCCACAGCTGGATTGGCAGCCCTTCTGTTGGCCGTGGCTGCTGGTGTCTTACTGT
GAGCTTCTCCACCTGGAAATTTCCCTCTCACCTACTTCTCTGGCCCTGGGTACCCCTCTTCT
CATCACTTCCTGTTCCCACTGACTGGGCTGGCCCAGCCCCTGTTTTTCCAACATTCCC
CAGTATCCCCAGCTTCTGCTGCGCTGGTTTTCGGCTTTGGGAAATAAAATACCGTTGTATAT
ATTCTGCCAGGGGTGTTCTAGCTTTTTGAGGACAGCTCCTGTATCCTTCTCATCCTTGTCTC
TCCGCTTGTCCTCTTGTGATGTTAGGACAGAGTGAGAGAAGTCAGCTGTCACGGGGAAGGTG
AGAGAGAGGATGCTAAGCTTCCTACTCACTTCTCCTAGCCAGCCTGGACTTTGGAGCGTGG
GGTGGGTGGGACAATGGCTCCCCACTCTAAGCACTGCCTCCCCTACTCCCCGCATCTTTGGG
GAATCGGTTCCCCATATGTCTTCCTTACTAGACTGTGAGCTCCTCGAGGGGGGGCCCGGTAC
CCAATTCGCCCTATAGTGAGTCGTA

125/330

FIGURE 125

MDPARKAGAQAMIWTAGWLLLLLLRGGAQALECYSCVQKADDGCSPNKMKTVKCAPGVDVCT
EAVGAVETIHGQFSLAVRGCGLPGKNDRGLDLHGLLAFIQLQQCAQDRCAKLNLTSRAL
DPAGNESAYPPNGVECYSCVGLSREACQGTSPPVVSCYNASDHVYKGCDFGNVTTLTAAVTV
SLPVRGCVQDEFCTR DGVTGPGFTLSGSCCQGSRCNSDLRNKTYFSPRIPLVRLPPPEPTT
VASTTSVTTSTTSAPVRPTSTTKPMPAPTSQTFRQVEHEASRDEEPRLTGGAAGHQDRSMSG
QYPAKGGPQQPHNKGCVAPTAGLAALLLAVAAGVLL

FIGURE 126

CGGGACTCGGCGGGTCCTCCTGGGAGTCTCGGAGGGGACCGGCTGTGCAGACGCC**ATG**GAGT
TGGTGCTGGTCTTCCTCTGCAGCCTGCTGGCCCCCATGGTCTGGCCAGTGCAGCTGAAAAG
GAGAAGGAAATGGACCCTTTTCATTATGATTACCAGACCCTGAGGATTGGGGGACTGGTGTT
CGCTGTGGTCCTCTTCTCGGTTGGGATCCTCCTTATCCTAAGTCGCAGGTGCAAGTGCAGTT
TCAATCAGAAGCCCCGGGCCCCAGGAGATGAGGAAGCCCAGGTGGAGAACCTCATCACCGCC
AATGCAACAGAGCCCCAGAAGCAGAGAACTGAAGTGCAGCCATCAGGTGGAAGCCTCTGGAA
CCTGAGGCGGCTGCTTGAACCTTTGGATGCAAATGTCGATGCT**TAAG**AAAACCGGCCACTTC
AGCAACAGCCCTTTCCCCAGGAGAAGCCAAGAACTTGTGTGTCCCCACCCTATCCCCTCTA
ACACCATTTCCTCCACCTGATGATGCAACTAACACTTGCCTCCCCACTGCAGCCTGCGGTCTT
GCCACCTCCCGTGATGTGTGTGTGTGTGTGTGTGTGACTGTGTGTGTTTGCTAACTGTG
GTCTTTGTGGCTACTTGTTTGTGGATGGTATTGTGTTTGTAGTGAAGTGTGGACTCGCTTT
CCCAGGCAGGGGCTGAGCCACATGGCCATCTGCTCCTCCCTGCCCCCGTGGCCCTCCATCAC
CTTCTGCTCCTAGGAGGCTGCTTGTTGCCCGAGACCAGCCCCCTCCCCTGATTTAGGGATGC
GTAGGGTAAGAGCACGGGCAGTGGTCTTCAGTCGTCTTGGGACCTGGGAAGGTTTGCAGCAC
TTTGTGCATCATTCTTCATGGACTCCTTTCACTCCTTTAACAAAAACCTTGCTTCCTTATCCC
ACCTGATCCCAGTCTGAAGGTCTCTTAGCAACTGGAGATACAAAGCAAGGAGCTGGTGAGCC
CAGCGTTGACGTCAGGCAGGCTATGCCCTTCCGTGGTTAATTTCTTCCCAGGGGCTTCCACG
AGGAGTCCCCATCTGCCCCGCCCTTCACAGAGCGCCCGGGGATTCCAGGCCCAGGGCTTCT
ACTCTGCCCTGGGGAATGTGTCCCCTGCATATCTTCTCAGCAATAACTCCATGGGCTCTGG
GACCCTACCCCTTCCAACCTTCCCTGCTTCTGAGACTTCAATCTACAGCCCAGCTCATCCAG
ATGCAGACTACAGTCCCTGCAATTGGGTCTCTGGCAGGCAATAGTTGAAGGACTCCTGTTCC
GTTGGGGCCAGCACACCGGGATGGATGGAGGGAGAGCAGAGGCCTTTGCTTCTCTGCCTACG
TCCCCTTAGATGGGCAGCAGAGGCAACTCCCGCATCCTTTGCTCTGCCTGTCGGTGGTCAGA
GCGGTGAGCGAGGTGGGTGGAGACTCAGCAGGCTCCGTGCAGCCCTTGGGAACAGTGAGAG
GTTGAAGGTCATAACGAGAGTGGGAACTCAACCCAGATCCCGCCCCTCCTGTCTCTGTGT
CCCGCGGAAACCAACCAAACCGTGCGCTGTGACCCATTGCTGTTCTCTGTATCGTGATCTAT
CCTCAACAACAACAGAAAAAAGGAATAAAATATCCTTTGTTTCCT

127/330

FIGURE 127

MELVLVFLCSLLAPMVLASAAEKEKEMDPFHYDYQTLRIGGLVFAVVLFSVGILLILSRCK
CSFNQKPRAPGDEEAQVENLITANATEPQKQORTEVQPSGGSLWNLRRLLLEPLDANVDA

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 128

AAACTTGACGCCATGAAGATCCCGGTCCTTCCTGCCGTGGTGCTCCTCTCCCTCCTGGTGCT
CCTCTGCCCAGGGAGCCACCCTGGGTGGTCCTGAGGAAGAAAGCACCATTGAGAATTATG
CGTCACGACCCGAGGCCTTTAACACCCCGTTCCTGAACATCGACAAATTGCGATCTGCGTTT
AAGGCTGATGAGTTCCTGAACTGGCACGCCCTCTTTGAGTCTATCAAAGGAACTTCCTTT
CCTCAACTGGGATGCCTTTCCTAAGCTGAAAGGACTGAGGAGCGCAACTCCTGATGCCCAGT
GACCATGACCTCCACTGGAAGAGGGGGCTAGCGTGAGCGCTGATTCTCAACCTACCATAACT
CTTCCTGCCTCAGGAACCTCAATAAAACATTTTCCATCCAAA

128/330

129/330

FIGURE 129

MKIPVLPVAVVLLSLLVLHSAQGATLGGPEEEESTIENYASRPEAFNTPFLNIDKLRSFAKADE
FLNWHALFESIKRKLPFLNWDAFPCLKGLRSATPDAQ

129/330

FIGURE 130

CAGTTCTGAAATCAATGGAGTTAATTTAGGGAATACAAACCAGCCATGGGGGTGGAGATTGC
CTTTGCCTCAGTGATTCTCACCTGCCTCTCCCTTCTGGCAGCAGGAGTCTCCCAGGTTGTTC
TTCTCCAGCCAGTTCCAACCTCAGGAGACAGGTCCCAAGGCCATGGGAGATCTCTCCTGTGGC
TTTGCCGGCCACTCATGAGAGTGTTTTTGTGTAAAGTATTTTTTTAGAATACTGTTGACTTCT
TCATGATTTAATAACCATCCTTTGCGAAGTTTTATGAGGCTTTAGGGGAATGTCAACCCCTCA
AATTTTTGTTATACTAGATGGCTTCCATTTACCCACCACTATTTTAAGGTCCCTTTATTTTT
AGGTTCAAGGTTCAATTTGACTTGAGAAAGTGCCCTTCTGCAGCTTCATTGATTTTGTTCATC
TTCATATTAAATTGTAACGATTAAAAAGAATAAGAGCACGCAGACCTCTAGGAGAATATTT
TATCCCTGGGTGCCCCTGACACATTTATGTAGTGATCCACAAATGTGATTGTTAATTTAAA
TGTTATTCTAATATTAGTACATTCAGTTGTGATGTAATATGAATAACCAGAATCTATTTCTT
AAAAGTTTTGAGTATATTTTTCAACTAGATATTTGTATAGAAAGACTGAATAGTGATG

131/330

FIGURE 131

MGVEIAFASVILTCLSLAAGVSQVLLQPVPTQETGPKAMGDLSCGFAGHS

131/330

GGGGAATCTGCAGTAGGTCTGCCGCGC**ATG**GAGTGGTGGGCTAGCTCGCCGCTTCGGCTCTG
GCTGCTGTTGTTCCCTCCTGCCCTCAGCGCAGGGCCGCCAGAAGGAGTCAGGTTCAAATGGA
AAGTATTTATTGACCAAATTAACAGGTCTTTGGAGAATTACGAACCATGTTCAAGTCAAAC
TGCAGCTGCTACCATGGTGT CATAGAAGAGGATCTAACTCCTTTCCGAGGAGGCATCTCCAG
GAAGATGATGGCAGAGGTAGTCAGACGGAAGCTAGGGACCCACTATCAGATCACTAAGAACA
GACTGTACCGGGAAAATGACTGCATGTTCCCCTCAAGGTGTAGTGGTGTGAGCACTTTATT
TTGGAAGTGATCGGGCGTCTCCCTGACATGGAGATGGTGATCAATGTACGAGATTATCCTCA
GGTTCCTAAATGGATGGAGCCTGCCATCCCAGTCTTCTCCTTCAGTAAGACATCAGAGTACC
ATGATATCATGTATCCTGCTTGGACATTTTGGGAAGGGGGACCTGCTGTTTGGCCAATTTAT
CCTACAGGTCTTGGACGGTGGGACCTCTTCAGAGAAGATCTGGTAAGGT CAGCAGCACAGTG
GCCATGGAAAAAGAAAACTCTACAGCATATTTCCGAGGATCAAGGACAAGTCCAGAACGAG
ATCCTCTCATTCTTCTGTCTCGGAAAAACCCAAAAC TTTGTTGATGCAGAATACACCAAAAAC
CAGGCCTGGAAATCTATGAAAGATACCTTAGGAAAGCCAGCTGCTAAGGATGTCCATCTTGT
GGATCACTGCAAATACAAGTATCTGTTTAATTTTCGAGGCGTAGCTGCAAGTTTCCGGTTTA
AACACCTCTTCTGTGTGGCTCACTTGTTTTCCATGTTGGTGATGAGTGGCTAGAATTCTTC
TATCCACAGCTGAAGCCATGGGTTC ACTATATCCAGTCAAACAGATCTCTCCAATGTCCA
AGAGCTGTTACAATTTGTAAAAGCAAATGATGATGTAGCTCAAGAGATTGCTGAAAGGGGAA
GCCAGTTTATTAGGAACCATTTGCAGATGGATGACATCACCTGTTACTGGGAGAACCTCTTG
AGTGAATACTCTAAATTCCTGTCTTATAATGTAACGAGAAGGAAAGGTTATGATCAAATTAT
TCCCAAATGTTGAAAAC TGAAC T**AG**TAGTCATCATAGGACCATAGTCCTCTTTGTGGCA
ACAGATCTCAGATATCCTACGGTGAGAAGCTTACCATAAGCTTGGCTCCTATACCTTGAATA
TCTGCTATCAAGCCAAATACCTGGTTTTCTTATCATGCTGCACCCAGAGCAACTCTTGAGA
AAGATTTAAAATGTGTCTAATACTGATATGAAGCAGTTCAACTTTTTGGATGAATAAGGA
CCAGAAATCGTGAGATGTGGATTTTGAACCCAACTCTACCTTTCAATTTCTTAAGACCAATC
ACAGCTTG TGCCCTCAGATCATCCACCTGTGTGAGTCCATCACTGTGAAATTGACTGTGTCCA
TGTGATGATGCCCTTTGTCCCAT TATTTGGAGCAGAAAATTCGTCAATTTGGAAGTAGTACAA
CTCATTGCTGGAATTGTGAAATTATTC AAGGCGTGATCTCTGTCACTTTATTTTAATGTAGG
AAACCCTATGGGGTTTTATGAAAAATACTTGGGGATCATTCTCTGAATGGTCTAAGGAAGCGG
TAGCCATGCCATGCAATGATGTAGGAGTTCTCTTTTGTA AAACCATAAACTCTGTTACTCAG
GAGGTTTCTATAATGCCACATAGAAAGAGGCCAATTGCATGAGTAATTATTGCAATTGGATT
TCAGGTTCCCTTTTTTGTGCCTTCATGCCCTACTTCTTAATGCCTCTCTAAAGCCAAA

FIGURE 133

MEWWASSPLRLWLLLFLLP
SAQGRQKESGSKWKVFIDQ
INRSLENYEPCSSQNCSCY
HGVIE
EDLTPFRGGISRKMMAEV
VRRKLGTHYQITKNRLYRE
NDCMFPSRCSGVEHFILEV
IGRLPD
MEMVINVRDYPQVPKWME
PAIPVFSFSKTSEYHDIMY
PAWTFWEGGPAVWPPIYPT
GLGRWDL
FREDLVRSAQWPWKKKNST
AYFRGSRTSPERDPLILLS
SRKNPKLVDAEYTKNQAWK
SMKDT
LGKPAAKDVHLVDHCKYK
YLENFRGVAASFRFKHLFL
CGSLVFHVGDEWLEFFYP
QLKPWVH
YIPVKTDLSNVQELLQFV
KANDDVAQEIAERGSQFIR
NHLQMDDITCYWENLLSEY
SKFLSY
NVTRRKGVDQIIPKMLK
TEL

133/330

FIGURE 134

CACCCCTCCATTTCTCGCC**ATG**GCCCCCTGCACTGCTCCTGATCCCTGCTGCCCTCGCCTCTT
 TCATCCTGGCCTTTGGCACCGGAGTGGAGTTCGTGCGCTTTACCTCCCTTCGGCCACTTCTT
 GGAGGGATCCCGGAGTCTGGTGGTCCGGATGCCCCGCCAGGGATGGCTGGCTGCCCTGCAGGA
 CCGCAGCATCCTTGCCCCCTGGCATGGGATCTGGGGCTCCTGCTTCTATTTGTTGGGCAGC
 ACAGCCTCATGGCAGCTGAAAGAGTGAAGGCATGGACATCCCGGTACTTTGGGGTCCTTCAG
 AGGTCACCTGTATGTGGCCTGCACTGCCCTGGCCTTGACAGCTGGTGATGCGGTACTGGGAGCC
 CATACCCAAAGGCCCTGTGTTGTGGGAGGCTCGGGCTGAGCCATGGGCCACCTGGGTGCCGC
 TCCTCTGCTTTGTGCTCCATGTCATCTCCTGGCTCCTCATCTTTAGCATCCTTCTCGTCTTT
 GACTATGCTGAGCTCATGGGCCTCAAACAGGTATACTACCATGTGCTGGGGCTGGGCGAGCC
 TCTGGCCCTGAAGTCTCCCCGGGCTCTCAGACTCTTCTCCACCTGCGCCACCCAGTGTGTG
 TGGAGCTGCTGACAGTGTGTGGGTGGTGCCTACCCTGGGCACGGACCGTCTCCTCCTTGCT
 TTCCTCCTTACCCTCTACCTGGGCCTGGCTCACGGGCTTGATCAGCAAGACCTCCGCTACCT
 CCGGGCCCAGCTACAAAGAAAACCTCCACCTGCTCTCTCGGCCCCAGGATGGGGAGGCAGAGT
GAGGAGCTCACTCTGGTTACAAGCCCTGTTCTTCTCCTCTCCCACTGAATTCTAAATCCTTAAC
 ATCCAGGCCCTGGCTGCTTCATGCCAGAGGCCCAAATCCATGGACTGAAGGAGATGCCCCCTT
 CTACTACTTGAGACTTTATTCTCTGGGTCCAGCTCCATACCCTAAATTCTGAGTTTCAGCCA
 CTGAACTCCAAGGTCCACTTCTCACCAGCAAGGAAGAGTGGGGTATGGAAGTCATCTGTCCC
 TTCCTGTTTGTAGAGCATGACACTCTCCCCCTCAACAGCCTCCTGAGAAGGAAAGGATCTGCC
 CTGACCACTCCCCCTGGCACTGTTACTTGCCCTCTGCGCCTCAGGGGTCCCCTTCTGCACCGCT
 GGCTTCCACTCCAAGAAGGTGGACCAGGGTCTGCAAGTTCAACGGTCATAGCTGTCCCTCCA
 GGCCCCAACCTTGCCCTCACCCTCCCGGCCCTAGTCTCTGCACCTCCTTAGGCCCTGCCTCT
 GGGCTCAGACCCCAACCTAGTCAAGGGGATTCTCCTGCTCTTAACTCGATGACTTGGGGCTC
 CCTGCTCTCCCGAGGAAGATGCTCTGCAGGAAAATAAAAGTCAGCCTTTTTCTAAAAAAA

FIGURE 135

MAPALLLIPAAALASFILAFGTGVEFVRFTSLRPLLGGIPESGGPDARQGWLAAALQDRSILAP
LAWDLGLLLLLFVGQHSILMAAERVKAWTSRYFGVLQRSLYVACTALALQLVMRYWEPIPKGPV
LWEARAEPWATWVPLLCFVLHVISWLLIFSILLVFDYAELMGLKQVYYHVLGLGEPLALKSP
RALRLFSLRLRHPVCVELLTVLWVPTLGTDRLLLAFLLTLYLGLAHGLDQQDLRYLRAQLQR
KLHLLSRPQDGEAE

Signal sequence:

amino acids 1-13

Transmembrane domains:

amino acids 58-76, 99-113, 141-159, 203-222

N-myristoylation sites:

amino acids 37-43, 42-48, 229-235

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 136

CCGAGCACAGGAGATTGCCTGCGTTTAGGAGGTGGCTGCGTTGTGGGAAAAGCTATCAAGGA
 AGAAATTGCCAAACCATGTCTTTTTTTCTGTTTTTCAGAGTAGTTCACAACAGATCTGAGTGT
 TTTAATTAAGCATGGAATACAGAAAACAACAAAAAAGCTTAAGCTTTAATTTTCATCTGGAATT
 CCACAGTTTTCTTAGCTCCCTGGACCCGGTTGACCTGTTGGCTCTTCCCGCTGGCTGCTCTA
 TCACGTGGTGTCTCCGACTACTCACCCCGAGTGTAAGAACCTTCGGCTCGCGTGCTTCTG
 AGCTGCTGTGGATGGCCTCGGCTCTCTGGACTGTCCTTCCGAGTAGGATGTCACTGAGATCC
 CTCAAATGGAGCCTCCTGCTGCTGTCACTCCTGAGTTTCTTTGTGATGTGGTACCTCAGCCT
 TCCCCACTACAATGTGATAGAACGCGTGAAGTGGATGTACTTCTATGAGTATGAGCCGATTT
 ACAGACAAGACTTTCACTTCACACTTCGAGAGCATTCAAAGTCTCTCATCAAAATCCATTT
 CTGGTCATTCTGGTGACCTCCCACCCTTCAGATGTGAAAGCCAGGCAGGCCATTAGAGTTAC
 TTGGGGTGAAAAAAGTCTTGGTGGGGATATGAGGTTCTTACATTTTTCTTATTAGGCCAAG
 AGGCTGAAAAGGAAGACAAAATGTTGGCATTGTCCTTAGAGGATGAACACCTTCTTTATGGT
 GACATAATCCGACAAGATTTTTTAGACACATATAATAACCTGACCTTGAAAACCATTTATGGC
 ATTCAGGTGGGTAAGTGAAGTTTTGCCCAATGCCAAGTACGTAATGAAGACAGACACTGATG
 TTTTCATCAATACTGGCAATTTAGTGAAGTATCTTTTAAACCTAAACCACTCAGAGAAGTTT
 TTCACAGGTATCCTCTAATTGATAATTATTCCTATAGAGGATTTTACCAAAAAACCCATAT
 TTCTTACCAGGAGTATCCTTTCAAGGTGTTCCCTCCATACTGCAGTGGGTTGGGTTATATAA
 TGTCCAGAGATTTGGTGCCAAGGATCTATGAAATGATGGGTCACGTAAAACCCATCAAGTTT
 GAAGATGTTTATGTCGGGATCTGTTTGAATTTATTAAAAGTGAACATTCATATTCAGAAGA
 CACAAATCTTTTCTTTCTATATAGAATCCATTTGGATGTCTGTCAACTGAGACGTGTGATTG
 CAGCCCATGGCTTTTCTTCCAAGGAGATCATCACTTTTTGGCAGGTCATGCTAAGGAACACC
 ACATGCCATTATTAAGCTTCACATTCTACAAAAGCCTAGAAGGACAGGATACCTTGTGGAAA
 GTGTTAAATAAAGTAGGTACTGTGAAAATTCATGGGGAGGTCAGTGTGCTGGCTTACACTG
 AACTGAAACTCATGAAAACCCAGACTGGAGACTGGAGGGTTACACTTGTGATTTATTAGTC
 AGGCCCTTCAAAGATGATATGTGGAGGAATTAAATATAAAGGAATTGGAGGTTTTTGCTAAA
 GAAATTAATAGGACCAAACAATTTGGACATGTCATTCTGTAGACTAGAATTTCTTAAAAGGG
 TGTTACTGAGTTATAAGCTCACTAGGCTGTAAAAACAAAACAATGTAGAGTTTTATTATTG
 AACAATGTAGTCACTTGAAGGTTTTGTGTATATCTTATGTGGATTACCAATTTAAAAATATA
 TGTAGTTCTGTGTCAAAAAACTTCTTCACTGAAGTTATACTGAACAAAATTTTACCTGTTTT
 TGGTCATTTATAAAGTACTTCAAGATGTTGCAGTATTTACAGTTATTATTATTAAAATTA
 CTTCAACTTTGTGTTTTTAAATGTTTTGACGATTTCAATACAAGATAAAAAGGATAGTGAAT
 CATTCTTTACATGCAAACATTTTCCAGTACTTAACTGATCAGTTTATTATTGATACATCAC
 TCCATTAATGTAAAGTCATAGGTCATTATTGCATATCAGTAATCTCTTGGACTTTGTAAAT
 ATTTTACTGTGGTAATATAGAGAAGAATTAAAGCAAGAAAATCTGAAAA

FIGURE 137

MASALWTVLPSRMSLRSLKWSLLLLSLLSFFVMWYLSLPHYNVIERVNWMYFYEYEPiYRQD
FHFTLREHSNC SHQNPFLVILVTSHPSDVKARQAIRVTWGEKKSWWGYEVLTFLLGQEA EK
EDKMLALSLEDEHLLYGDIIRQDFLDTYNNLTLKTIMAFRWVTEFCPNAKYVMKTD TDVFIN
TGNLVKYLLNLNHSEKFFTGYPLIDNYSYRGFYQKTHISYQEYPFKVFP PYCSGLGYIMSRD
LVPRIYEMMGHV KPIKFEDVYVGICLNLLKVNIHIPEDTNLFFLYRIHLDVCQLRRVIAAHG
FSSKEIITFWQVMLRNTTCHY

137/330

FIGURE 138

CCTCTGTCCACTGCTTTCGTGAAGACAAGATGAAGTTCACAATTGTCTTTGCTGGACTTCTT
GGAGTCTTTCTAGCTCCTGCCCTAGCTAACTATAATATCAACGTCAATGATGACAACAACAA
TGCTGGAAGTGGGCAGCAGTCAGTGAGTGTCAACAATGAACACAATGTGGCCAATGTTGACA
ATAACAACGGATGGGACTCCTGGAATTCCATCTGGGATTATGGAAATGGCTTTGCTGCAACC
AGACTCTTTCAAAAGAAGACATGCATTGTGCACAAAATGAACAAGGAAGTCATGCCCTCCAT
TCAATCCCTTGATGCACTGGTCAAGGAAAAGAAGCTTCAGGGTAAGGGACCAGGAGGACCAC
CTCCCAAGGGCCTGATGTACTCAGTCAACCCAAACAAAGTCGATGACCTGAGCAAGTTCGGA
AAAAACATTGCAAACATGTGTCGTGGGATTCCAACATACATGGCTGAGGAGATGCAAGAGGC
AAGCCTGTTTTTTTTACTCAGGAACGTGCTACACGACCAGTGTACTATGGATTGTGGACATTT
CCTTCTGTGGAGACACGGTGGAGAACTTAAACAATTTTTTAAAGCCACTATGGATTTAGTCAT
CTGAATATGCTGTGCAGAAAAAATATGGGCTCCAGTGGTTTTTACCATGTCATTCTGAAATT
TTTCTCTACTAGTTATGTTTGATTTCCTTAAGTTTCAATAAAATCATTTAGCATTGAAAAAAA

FIGURE 139

MKFTIVFAGLLGVFLAPALANYNINVNDDNNNAGSGQQSVSVNNEHNVANVDNNNGWDSWNS
IWDYGNGFAATRLFQKKTCIVHKMNKEVMPSIQSLDALVKEKKLQKGPGGPPPKGLMYSVN
PNKVDDLKFGKNIANMCRGIPTYMAEEMQEASLFFYSGTCYTTSVLWIWDISFCGDTVEN

Signal Peptide:

amino acids 1-20

N-myristoylation Sites:

amino acids 67-72, 118-123, 163-168

Flavodoxin protein homology:

amino acids 156-174

139/330

141/330

FIGURE 141

MNDSLRTNVFVRFQPETIACACIYLAARALQIPLPTRPHWFLLFGTTEEEIQEICIETLRLY
TRKKPNYELLEKEVEKRKVALQEAKLKAKGLNPDGTPALSTLGGFSPASKPSSPREVKAEK
SPISINVKTVKKEPEDRQQASKSPYNGVRKDSKRSRNSRSASRSRSTRSRSRSHTPRRHYN
NRRSRSGTYSSRSRSHSESPRRHHNHGSPHLKAKHTRDDLKSSNRHGHKRRKRSRSQ
SKSRDHSDAAKKHRHERGHRDRRERSRSFERSHKS KHHGGSRSRSGHGRHRR

141/330

FIGURE 142

TGGGGATAAAGGAAAAATGGTCAGGTATTAATGGCTTAAAGATTATTGGAAGGGGTTTATCA
TTTTTTGAANNTATTCGGGTCANAATTGNCTTTGAAAAGCATTGCTTTTTACAGAAATATAT
TANCTTTTTAGAGTAATTTCTAGTTTGGATTGTAATATGAAATTATTTAAAAGGGCTTCGCT
CATATATAGGAAAATCGCATATGGTCCTAGTATTAATTTNTTATTGCTTACTGATTTTTTTG
AGTTAAGAGTTGTTATATGNTAGAATATGAGGATGTGAATATAAATAAGAGAAGAAAAAAGA
ATAAAGTAGATTGAGTCTCCAATTTTATGTAAGCTTCAGAAGAAGTGGTTTGTTTACATGCA
AGCTTATAGTTGAAATATTTTTTCAGGAATTACATGAATGACAGTCTTCGAACCAATGTGTTT
GTTTCGATTTCAACCAGAGANTATAGCATGTGCTTGCACTACCTTGCAGNTAGAGCACTTCA
GATTCCGTTGCCAACTNGTCCCCATTGGTTTCTTCTTTTTGGTACTACAGAAGAGGAAATCC
AGGAAATNTGCATAGAAACACTTAGGCTTTATACCAGAAAAAAGCCAACTATGAATTACTG
GAAAAAGAAGTAGAAAAAAGAAAAGTAGCCTTACAAGAAGCCNAATTAAAAGCAAAGGGATT
GAATCCGGATGGAAGTCCAGCCCTTTCAACCCTGGGTGGATTTTCTCC

FIGURE 143

GGCACGAGGCCTCGTGCCAAGCTTGGCACGAGGGTGCACCGCGTTCTCGCACGCGT**CATGGC**
GGTCCTCGGAGTACAGCTGGTGGTGACCCTGCTCACTGCCACCCTCATGCACAGGCTGGCGC
CACACTGCTCCTTCGCGCGCTGGCTGCTCTGTAACGGCAGTTTGTTCGATACAAGCACCCG
TCTGAGGAGGAGCTTCGGGGCCCTGGCGGGGAAGCCGAGGCCAGAGGCAGGAAAGAGCGGTG
GGCCAATGGCCTTAGTGAGGAGAAGCCACTGTCTGTGCCCCGAGATGCCCCGTTCCAGCTGG
AGACCTGCCCCCTCACGACCGTGGATGCCCTGGTCCTGCGCTTCTTCCTGGAGTACCAGTGG
TTTGTGGACTTTGCTGTGTACTCGGGCGGCGTGTACCTCTTCACAGAGGCCTACTACTACAT
GCTGGGACCAGCCAAGGAGACTAACATTGCTGTGTTCTGGTGCCTGCTCACGGTGACCTTCT
CCATCAAGATGTTCTGACAGTGACACGGCTGTACTTCAGCGCCGAGGAGGGGGGTGAGCGC
TCTGTCTGCCTCACCTTTGCCTTCCTCTTCCTGCTGCTGGCCATGCTGGTGCAAGTGGTGCG
GGAGGAGACCTCGAGCTGGGCCTGGAGCCTGGTCTGGCCAGCATGACCCAGAACTTAGAGC
CACTTCTGAAGAAGCAGGGCTGGGACTGGGCGCTTCCTGTGGCCAAGCTGGCTATCCGCGTG
GGACTGGCAGTGGTGGGCTCTGTGCTGGGTGCCTTCCTCACCTTCCCAGGCCTGCGGCTGGC
CCAGACCCACCGGGACGCACTGACCATGTGCGAGGACAGACCCATGCTGCAGTTCCTCCTGC
ACACCAGCTTCCTGTCTCCCCTGTTTCATCCTGTGGCTCTGGACAAAGCCCATTGCACGGGAC
TTCCTGCACCAGCCGCCGTTTGGGGAGACGCGTTTCTCCCTGCTGTCCGATTCTGCCTTCGA
CTCTGGGCGCCTCTGGTTGCTGGTGGTGTGCTGTGCCTGCTGCGGCTGGCGGTGACCCGGCCCC
ACCTGCAGGCCTACCTGTGCCTGGCCAAGGCCCGGGTGGAGCAGCTGCGAAGGGAGGCTGGC
CGCATCGAAGCCCGTGAAATCCAGCAGAGGGTGGTCCGAGTCTACTGCTATGTGACCGTGGT
GAGCTTGCAGTACCTGACGCCGCTCATCCTCACCTCAACTGCACACTTCTGCTCAAGACGC
TGGGAGGCTATTCTGGGGCCTGGGCCCAGCTCCTCTACTATCCCCGACCCATCCTCAGCC
AGCGCTGCCCCCATCGGCTCTGGGGAGGACGAAGTCCAGCAGACTGCAGCGCGGATTGCCGG
GGCCCTGGGTGGCCTGCTTACTCCCCTCTTCCTCCGTGGCGTCCTGGCCTACCTCATCTGGT
GGACGGCTGCCTGCCAGCTGCTCGCCAGCCTTTTCGGCCTCTACTTCCACCAGCACTTGGCA
GGCTCC**TAG**CTGCCTGCAGACCCTCCTGGGGCCCTGAGGTCTGTTCTGGGGCAGCGGGACA
CTAGCCTGCCCCCTCTGTTTGCGCCCCCGTGTCCCCAGCTGCAAGGTGGGGCCGGACTCCCC
GGCGTTCCCTTCACCACAGTGCCTGACCCGCGGCCCCCCTTGGACGCCGAGTTTCTGCCTCA
GAACTGTCTCTCCTGGGCCCAGCAGCATGAGGGTCCCGAGGCCATTGTCTCCGAAGCGTATG
TGCCAGGTTTGAGTGGCGAGGGTGATGCTGGCTGCTCTTCTGAACAAATAAAGGAGCATGCC
GATTTTAA

FIGURE 144

MAVLGVQLVVTLLTATLMHRLAPHCSFARWLLCNGSLFRYKHPSEEELRALAGKPRPRGRKE
RWANGLSEEKPLSVPRDAPFQLETCPLTTVDALVLRFFLEYQWFVDFAVYSGGVYLFTEAYY
YMLGPAKETNIAVFWCLLTVTFSIKMFLTVTRLYFSAEEGGERSVCLTFAFLFLLLAMLVQV
VREETLELGLEPGLASMTQNLEPLLKKQGWDWALPVAKLAIKRVGLAVVGSVLGAFLTFPGLR
LAQTHRDALTMSEDRPMLQFLLHTSFLSPLFILWLWTKPIARDFLHQPFFGETRFSLLSDSA
FDSGRLWLLVVLCLLRLAVTRPHLQAYLCLAKARVEQLRREAGRIEAREIQQRVVRVYCYVT
VVSLLQYLTPLIILTINCTLLLKTLGGYSWGLGPAPLLSPDPSSASAAPIGSGEDEVQOTAARI
AGALGGLLTPLFLRGVLAYLIWWTAACQLLASLFGLYFHQHLA

FIGURE 145

CGTTNGCACGCGTCAATGGCGGTCCTCGGAGTACAGCTGGTGGTGACCCTGCTCACTGCCAC
CCTCATGCACAGGCTGGCGCCACACTGCTCCTTCGCGCGCTGGCTGCTCTGTAAACGGCAGTT
TGTTCCGATACAAGCACCCGTNTTGAGGAGGAGCTTCGGGCCCTGGCGGGGAAGCCGAGGCC
CAGAGGCAGGAAAGAGCGGTGGGCCAATGGCCTTAGTGAGGAGAAGCCACTGTCTGTGCCCC
GAGATGCCCCGTTCAGCTGGAGACCTGCCCCCTCACGACCGTGGATGCCCTGGTCCTGCGC
TTCTTCCTGGAGTACCAGTGGTTTGTGGACTTTGCTGTGTACTCGGGCGGCGTGTACCTCTT
CACAGAGGCCTACTACTACATGCTGGGACCAGCCAAGGAGACTAACATTGCTGTGTTCTGGT
GCCTGCTCACAGTGACCTTCTCCATCAAGATGTTTCCTGACAGTGACACGGCTGTACTTCAGC
GCCGAGGAGGGGGGTGAGCGCTCTGTCTGCCTCACCTTGCCTTCCTCTTCCTGCTGCTGGC
CATGCTGGTGCAAGCG

CGTTNGCACGCGTCAATGGCGGTCCTCGGAGTACAGCTGGTGGTGACCCTGCTCACTGCCAC
CCTCATGCACAGGCTGGCGCCACACTGCTCCTTCGCGCGCTGGCTGCTCTGTAAACGGCAGTT
TGTTCCGATACAAGCACCCGTNTTGAGGAGGAGCTTCGGGCCCTGGCGGGGAAGCCGAGGCC
CAGAGGCAGGAAAGAGCGGTGGGCCAATGGCCTTAGTGAGGAGAAGCCACTGTCTGTGCCCC
GAGATGCCCCGTTCAGCTGGAGACCTGCCCCCTCACGACCGTGGATGCCCTGGTCCTGCGC
TTCTTCCTGGAGTACCAGTGGTTTGTGGACTTTGCTGTGTACTCGGGCGGCGTGTACCTCTT
CACAGAGGCCTACTACTACATGCTGGGACCAGCCAAGGAGACTAACATTGCTGTGTTCTGGT
GCCTGCTCACAGTGACCTTCTCCATCAAGATGTTTCCTGACAGTGACACGGCTGTACTTCAGC
GCCGAGGAGGGGGGTGAGCGCTCTGTCTGCCTCACCTTGCCTTCCTCTTCCTGCTGCTGGC
CATGCTGGTGCAAGCG

FIGURE 146

GGTTCTACATCCTCTCATCTGAGAATCAGAGAGCATAATCTTCTTACGGGCCCGTGATTTATTAACGTGGCTT
AATCTGAAGGTTCTCAGTCAAATTCCTTGTGATCTACTGATTGTGGGGGCATGGCAAGGTTTGCTTAAAGGAGC
TTGGCTGGTTTGGGCCCTTGTAGCTGACAGAAGGTGGCCAGGGAGAATGCAGCACACTGCTCGGAGAATGAAGG
CGCTTCTGTTGCTGGTCTTGCCCTTGGCTCAGTCTGCTAACTACATTGACAATGTGGGCAACCTGCACTTCCTG
TATTCAGAACTCTGTAAAGGTGCCTCCCCTACGGCCTGACCAAAGATAGGAAGAGGCGCTCACAAGATGGCTG
TCCAGACGGCTGTGCGAGCCTCACAGCCACGGCTCCCTCCCCAGAGGTTTCTGCAGCTGCCACCATCTCCTTAA
TGACAGACGAGCCTGGCCTAGACAACCTGCCTACGTGTCTCGGCAGAGGACGGGCAGCCAGCAATCAGCCCA
GTGGACTCTGGCCGGAGCAACCGAACTAGGGCACGGCCCTTTGAGAGATCCACTATTAGAAGCAGATCATTTAA
AAAAATAAATCGAGCTTTGAGTGTCTTCGAAGGACAAAGAGCGGGAGTGAGTTGCCAACCATGCCGACCAGG
GCAGGGAAAATTCTGAAAACACCACTGCCCTGAAGTCTTTCCAAGGTTGTACCACCTGATTCCAGATGGTGAA
ATTACCAGCATCAAGATCAATCGAGTAGATCCCAGTGAAAGCCTCTCTATTAGGCTGGTGGGAGGTAGCGAAAC
CCCCTGGTCCATATCATTTATCCAACACATTTATCGTGATGGGGTGATCGCCAGAGACGGCCGGCTACTGCCAG
GAGACATCATTCTAAAGGTCAACGGGATGGACATCAGCAATGTCCCTCACAACCTACGCTGTGCGTCTCCTGCGG
CAGCCCTGCCAGGTGCTGTGGCTGACTGTGATGCGTGAACAGAAGTTCCGCAGCAGGAACAATGGACAGGCCCC
GGATGCCTACAGACCCCGAGATGACAGCTTTTATGTGATTCTCAACAAAAGTAGCCCCGAGGAGCAGCTTGGA
TAAAACTGGTGCGCAAGGTGGATGAGCCTGGGGTTTTTATCTTCAATGTGCTGGATGGCGGTGTGGCATATCGA
CATGGTCAGCTTGAGGAGAATGACCGTGTGTTAGCCATCAATGGACATGATCTTCGATATGGCAGCCAGAAAG
TGCGGCTCATCTGATTCAGGCCAGTGAAAGACGTGTTCACCTCGTGTCCCGCCAGGTTCCGCAGCGGAGCC
CTGACATCTTTCAGGAAGCCGGCTGGAACAGCAATGGCAGCTGGTCCCCAGGGCCAGGGGAGAGGAGCAACACT
CCCAAGCCCCCTCCATCCTACAATTACTTGTGATGAGAAGGTGGTAAATATCCAAAAGACCCCGGTGAATCTCT
CGGCATGACCGTCCGAGGGGGAGCATCACATAGAGAATGGGATTTGCCTATCTATGTCATCAGTGTTGAGCCCG
GAGGAGTCATAAGCAGAGATGGAAGAATAAAAAACAGGTGACATTTTGTGAAATGTGGATGGGGTCGAACTGACA
GAGGTGAGCCGGAGTGAGGCAGTGGCATTATTGAAAAGAACATCATCCTCGATAGTACTCAAAGCTTTGGAAGT
CAAAGAGTATGAGCCCCAGGAAGACTGCAGCAGCCAGCAGCCCTGGACTCCAACCACAACATGGCCCCACCCA
GTGACTGGTCCCCATCCTGGGTGATGTGGCTGGAATTACCACGGTGCTTGTATAACTGTAAAGATATTGTATTA
CGAAGAAACACAGCTGGAAGTCTGGGCTTCTGCATTGTAGGAGTTATGAAGAATACAATGGAAACAAACCTTT
TTTCATCAAATCCATTGTTGAAGGAACACCAGCATACAATGATGGAAGAATTAGATGTGGTGATATTCTTCTTG
CTGTCAATGGTAGAAGTACATCAGGAATGATACATGCTTGGTGGCAAGACTGCTGAAAGAACTTAAAGGAAGA
ATTACTCTAACTATTGTTTCTTGGCCTGGCACTTTTTTATAGAAATCAATGATGGGTGAGAGGAAAACAGAAAAA
TCACAAATAGGCTAAGAAGTTGAAACACTATATTTATCTTGTGAGTTTTTATATTTAAAGAAAGAATACATTGT
AAAAATGTCAGGAAAAGTATGATCATCTAATGAAAGCCAGTTACACCTCAGAAAATATGATTCCAAAAAATTA
AACTACTAGTTTTTTTTTTCAGTGTGGAGGATTTCTCATTACTCTACAACATTGTTTATATTTTTCTATTCAAT
AAAAAGCCCTAAAACAACATAAATGATTGATTGTATACCCCACTGAATTCAGCTGATTTAAATTTAAATTT
GGTATATGCTGAAGTCTGCCAAGGGTACATTATGGCCATTTTAAATTTACAGCTAAAATATTTTTTAAATGCA
TTGCTGAGAAACGTTGCTTTCATCAAACAAGAATAAATATTTTTTCAAGTTAAA

FIGURE 147

MKALLLLVLPWLSPANYIDNVGNLHFLYSELCKGASHYGLTKDRKRRSQDGCPDGCASLTAT
APSPEVSAAATISLMTDEPGLDNPAYVSSAEDGQPAISPVDSGRSNRTRARPFERSTIRSRS
FKKINRALSVLRRTKSGSAVANHADQGRENSENTTAPEVFFRPLYHLIPDGEITSIKINRVDP
SESLSIRLVGGSETPLVHIIIIQHIYRDGVIARDGRLLPGDIILKVNGMDISNVPHNYAVRLL
RQPCQVLWLTVMREQKFRSRNNGQAPDAYRPRDDSFHVILNKSSPEEQLGIKLVRKVDEPGV
FIFNVLDGGVAYRHGQLEENDRVLAINGHDLRYGSPESA AHLIQASERRVHLVVSQRQVRS
PDIFQEAGWNSNGSWSPGPGERSNTPKPLHPTITCHEKVVNIQKDPGESLGMTVAGGASHRE
WDLPIYVISVEPGGVISR DGRIKTGDILLNVDGVELTEVSRSEAVALLKRTSSSIVLKALEV
KEYEPQEDCSSPAALDSNHNMAPPSDWSPSWVMWLELPRCLYNCKDIVLRRNTAGSLGFCIV
GGYEEYNGNKPFFIKSIVEGTPAYNDGRIRCGDILLAVNGRSTSGMIHACLARLLKELKGRI
TLTIVSWPGTFL

bioRxiv preprint doi: <https://doi.org/10.1101/147330>; this version posted May 1, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 148

CCAAAGTGATCATTTGAAAAAGAGATATCCACATCTTCAAGCCCATATAAAGGATAGAAGCT
GCACAGGGCAGCTTTACTTACTCCAGCACCTTCCTCTCCCAGGCAA**ATG**GTGCTGACCATCT
TTGGGATACAATCTCATGGATACGAGGTTTTTAACATCATCAGCCCAAGCAACAATGGTGGC
AATGTTCAGGAGACAGTGACAATTGATAATGAAAAAATACCGCCATCGTTAACATCCATGC
AGGATCATGCTCTTCTACCACAATTTTTGACTATAAACATGGCTACATTGCATCCAGGGTGC
TCTCCCGAAGAGCCTGCTTTATCCTGAAGATGGACCATCAGAACATCCCTCCTCTGAACAAT
CTCCAATGGTACATCTATGAGAAACAGGCTCTGGACAACATGTTCTCCAACAAATACACCTG
GGTCAAGTACAACCCCTCTGGAGTCTCTGATCAAAGACGTGGATTGGTTCCTGCTTGGGTCAC
CCATTGAGAACTCTGCAAACATATCCCTTTGTATAAGGGGGAAGTGGTTGAAAACACACAT
AATGTCGGTGCTGGAGGCTGTGCAAAGGCTGGGCTCCTGGGCATCTTGGGAATTTCAATCTG
TGCAGACATTCATGTT**TAG**GATGATTAGCCCTCTTGTTTTATCTTTTCAAAGAAATACATCC
TTGGTTTACACTCAAAGTCAAATTAAATTCTTTCCCAATGCCCCAACTAATTTTGAGATTC
AGTCAGAAAATATAAATGCTGTATTTATA

149/330

FIGURE 149

MKILVAFLVVLTIFGIQSHGYEVFNIISPSNNGGNVQETVTIDNEKNTAIVNIHAGSCSSTT
IFDYKHGYIASRVLSRRACFILKMDHQNIPPLNNLQWYIYEKQALDNMFSNKYTWVKYNPLE
SLIKDVDWFLLGSPIEKLCKHIPLYKGEVVENTHNVGAGGCAKAGLLGILGISICADIHV

149/330

FIGURE 150

GGCACGAGCCAGGAACTAGGAGGTTCTCACTGCCCCGAGCAGAGGCCCTACACCCACCGAGGC
ATGGGGCTCCCTGGGCTGTTCTGCTTGGCCGTGCTGGCTGCCAGCAGCTTCTCCAAGGCACG
GGAGGAAGAAATTACCCCTGTGGTCTCCATTGCCTACAAAGTCCTGGAAGTTTCCCCAAAG
GCCGCTGGGTGCTCATAACCTGCTGTGCACCCAGCCACCACCGCCCATCACCTATTCCCTC
TGTGGAACCAAGAACATCAAGGTGGCCAAGAAGGTGGTGAAGACCCACGAGCCGGCCTCCTT
CAACCTCAACGTCACACTCAAGTCCAGTCCAGACCTGCTCACCTACTTCTGCCGGGCGTCCT
CCACCTCAGGTGCCCATGTGGACAGTGCCAGGCTACAGATGCACTGGGAGCTGTGGTCCAAG
CCAGTGTCTGAGCTGCGGGCCAACCTTCACTCTGCAGGACAGAGGGGCAGGCCCCAGGGTGGA
GATGATCTGCCAGGCGTCCTCGGGCAGCCCACCTATCACCAACAGCCTGATCGGGAAGGATG
GGCAGGTCCACCTGCAGCAGAGACCATGCCACAGGCAGCCTGCCAACTTCTCCTTCCTGCCG
AGCCAGACATCGGACTGGTTCTGGTGCCAGGCTGCAAACAACGCCAATGTCCAGCACAGCGC
CCTCACAGTGGTGCCCCCAGGTGGTGACCAGAAGATGGAGGACTGGCAGGGTCCCCTGGAGA
GCCCCATCCTTGCCTTGCCGCTCTACAGGAGCACCCGCCGTCTGAGTGAAGAGGAGTTTGGG
GGGTTCAGGATAGGGAATGGGGAGGTCAGAGGACGCAAAGCAGCAGCCATG**TAGA**ATGAACC
GTCCAGAGAGCCAAGCACGGCAGAGGACTGCAGGCCATCAGCGTGCACTGTTTCGTATTTGGA
GTTTCATGCAAAATGAGTGTGTTTTAGCTGCTCTTGCCACAAAAAAAAAAAAAAAAAAAAA

151/330

FIGURE 151

MGLPGLFCLAVLAASSFSKAREEEITPVVSIAYKVLEVPKGRWVLITCCAPQPPPPITYSL
CGTKNIKVAKKVVKTHEPASFNLVTLKSSPDLLTYFCRASSTSGAHVDSARLQMHWELWSK
PVSELRANFTLQDRGAGPRVEMICQASSGSPITNSLIGKDGQVHLQQRPCHRQPANFSFLP
SQTSDWFWCQAANNANVQHSALTVVPPGGDQKMEDWQGPLESPILALPLYRSTRRLSEEEFG
GFRIGNGEVRGRKAAAM

Signal Peptide:

amino acids 1-18

N-glycosylation Sites:

amino acids 86-89, 132-135, 181-184

FIGURE 152

GGTCCTTAATGGCAGCAGCCGCCGCTACCAAGATCCTTCTGTGCCTCCCGCTTCTGCTCCTG
CTGTCCGGCTGGTCCCGGGCTGGGCGAGCCGACCCTCACTCTCTTTGCTATGACATCACCGT
CATCCCTAAGTTCAGACCTGGACCACGGTGGTGTGCGGTTCAAGGCCAGGTGGATGAAAAGA
CTTTTCTTCACTATGACTGTGGCAACAAGACAGTCACACCTGTCAGTCCCCTGGGGAAGAAA
CTAAATGTCACAACGGCCTGGAAAGCACAGAACCCAGTACTGAGAGAGGTGGTGGACATACT
TACAGAGCAACTGCGTGACATTAGCTGGAGAATTACACACCCAAGGAACCCCTCACCTGC
AGGCAAGGATGTCTTGTGAGCAGAAAGCTGAAGGACACAGCAGTGGATCTTGGCAGTTCAGT
TTCGATGGGCAGATCTTCCTCCTCTTTGACTCAGAGAAGAGAATGTGGACAACGGTTCATCC
TGGAGCCAGAAAGATGAAAGAAAAGTGGGAGAATGACAAGGTGTGGCCATGTCCTTCCATT
ACTTCTCAATGGGAGACTGTATAGGATGGCTTGAGGACTTCTTGATGGGCATGGACAGCACC
CTGGAGCCAAGTGCAGGAGCACCCTCGCCATGTCCTCAGGCACAACCCAACTCAGGGCCAC
AGCCACCACCCTCATCCTTTGCTGCCTCCTCATCATCCTCCCCTGCTTCATCCTCCCTGGCA
TCTGAGGAGAGTCCTTTAGAGTGACAGGTAAAGCTGATACCAAAGGCTCCTGTGAGCACG
GTCTTGATCAAACCTCGCCCTTCTGTCTGGCCAGCTGCCCACGACCTACGGTGTATGTCCAGT
GGCCTCCAGCAGATCATGATGACATCATGGACCCAATAGCTCATTCACTGCCTTGATTCCTT
TTGCCAACAATTTTACCAGCAGTTATACCTAACATATTATGCAATTTTCTCTTGGTGCTACC
TGATGGAATTCCTGCACTTAAAGTTCTGGCTGACTAAACAAGATATATCATTTTCTTTCTTC
TCTTTTGTGGAAAATCAAGTACTTCTTTGAATGATGATCTCTTTCTTGCAAATGATATT
GTCAGTAAATAATCACGTTAGACTTCAGACCTCTGGGGATTCTTTCCGTGTCCTGAAAGAG
AATTTTAAATTATTTAATAAGAAAAAATTTATATTAATGATTGTTTCCTTTAGTAATTTAT
TGTTCTGTACTGATATTTAAATAAAGAGTTCTATTTCCCAAAAAAAAAAAAAAAAAAAAA

FIGURE 153

MAAAAATKILLCLPLLLLLSGWSRAGRADPHSLCYDITVIPKFRPGPRWCAVQGQVDEKTFL
HYDCGNKTVTPVSPLGKKLNVTAWKAQNPVLREVVDILTEQLRDIQLENYTPKEPLTLQAR
MSCEQKAEGHSSGSWQFSFDGQIFLLFDSEKRMWTTVHPGARKMKEKWENDKVVAMSFHYFS
MGDCIGWLEDFLMGMDSTLEPSAGAPLAMSSGTTQLRATATTLILCCLLIILPCFILPGI

Important features:**Signal peptide:**

amino acids 1-25

Transmembrane domain:

amino acids 224-246

N-glycosylation site.

amino acids 68-72, 82-86

N-myristoylation site.

amino acids 200-206, 210-216

Amidation site.

amino acids 77-81

FIGURE 154

GGGAAAGCCATTTTCGAAAACCCATCTATACAACTATATATTTTCATTTCTGCTGCTAGCTG
CCTTGGGCCTCACAATTTTCATTCTGTTTTCTGACTTTCAAGTTATATACCGTGGA**ATG**GAG
TTGATCCCAACCATAACATCGTGGAGGGTTTTAATTTTGGTGGTAGCCCTCACCCAATTCTG
GTGTGGCTTTCTTTGCAGAGGATTCCACCTTCAAATCATGAACTCTGGCTGTTGATCAAAA
GAGAATTTGGATTCTACTCTAAAAGTCAATATAGGACTTGGCAAAGAAGCTAGCAGAAGAC
TCAACCTGGCCTCCCATAAACAGGACAGATTATTCAAGGTGATGGCAAAAATGGATTCTACAT
CAACGGAGGCTATGAAAGCCATGAACAGATTCCAAAAAGAAAATCAAATTGGGAGGCCAAC
CCACAGAACAGCATTTCTGGGCCAGGCTG**TAA**TCAGAATTGTCGTCGTACATGCTCAACAGC
ATTGCTTTTTTCCCCAAAATTAACACATTGTGGAGAAGTGATGATACTCTCCCCTTACCTTT
CCTCTCTCCATTCAAGCATTCAAAGTATATTTTCAATGAATTAAACCTTGCAGCAAGGGACC
TTAGATAGGCTTATTCTGACTGTATGCTTTACCAATGAGAGAAAAAAATGCATTTCCCTGTAT
CATCCTTTTCAATAAACTGTATTCATTTTGAAAAAAAAAAAAAAAAAAAAA

155/330

FIGURE 155

MELIPTITSWRVLILVVALTQFWCGFLCRGFHLQNHFWLLIKREFGFGYSSQYRTWQKLA
EDSTWPPINRTDYSGDGKNGFYINGGYESHEQIPKRKLKLGGQPTEQHFWARL

155/330

FIGURE 156

GTTCTCCTTTCCGAGCCAAAATCCCAGGCGATGGTGAATTATGAACGTGCCACACC**ATGAAG**
 CTCTTGTGGCAGGTAACTGTGCACCACCACACCTGGAATGCCATCCTGCTCCCGTTCGTCTA
 CCTCACGGCGCAAGTGTGGATTCTGTGTGCAGCCATCGCTGCTGCCGCCTCAGCCGGGCCCC
 AGAACTGCCCCCTCCGTTTGCTCGTGCAGTAACCAGTTCAGCAAGGTGGTGTGCACGCGCCGG
 GGCTCTCCGAGGTCCCGCAGGGTATTCCCTCGAACACCCGGTACCTCAACCTCATGGAGAA
 CAACATCCAGATGATCCAGGCCGACACCTTCCGCCACCTCCACCACCTGGAGGTCTGCAGT
 TGGGCAGGAACTCCATCCGGCAGATTGAGGTGGGGGCCTTCAACGGCCTGGCCAGCCTCAAC
 ACCCTGGAGCTGTTTCGACAACTGGCTGACAGTCATCCCTAGCGGGGCCTTTGAATACCTGTC
 CAAGCTGCGGGAGCTCTGGCTTCGCAACAACCCCATCGAAAGCATCCCCCTCTTACGCCTTCA
 ACCGGGTGCCCTCCCTCATGCGCCTGGACTTGGGGGAGCTCAAGAAGCTGGAGTATATCTCT
 GAGGGAGCTTTTGAGGGGCTGTTCAACCTCAAGTATCTGAACCTGGGCATGTGCAACATTAA
 AGACATGCCCAATCTCACCCCCCTGGTGGGGCTGGAGGAGCTGGAGATGTCAGGGAACCACT
 TCCCTGAGATCAGGCCTGGCTCCTTCCATGGCCTGAGCTCCCTCAAGAAGCTCTGGGTCTAG
 AACTCACAGGTCAGCCTGATTGAGCGGAATGCTTTTGACGGGCTGGCTTCACTTGTGGAAC
 CAACTTGGCCCAATAACCTCTCTTCTTTGCCCATGACCTCTTTACCCCGCTGAGGTACC
 TGGTGGAGTTGCATCTACACCACAACCCTTGGAACTGTGATTGTGACATTCTGTGGCTAGCC
 TGGTGGCTTCGAGAGTATATACCCACCAATTCCACCTGCTGTGGCCGCTGTCTGCTCCCAT
 GCACATGCGAGGCCGCTACCTCGTGGAGGTGGACCAGGCCTCCTTCCAGTGCTCTGCCCCCT
 TCATCATGGACGCACCTCGAGACCTCAACATTTCTGAGGGTCGGATGGCAGAACTTAAGTGT
 CGGACTCCCCCTATGTCCTCCGTGAAGTGGTTGCTGCCCAATGGGACAGTGCTCAGCCACGC
 CTCCCGCCACCAAGGATCTCTGTCTCAACGACGGCACCTTGAACCTTTTCCACGTGCTGC
 TTTTACAGACTGGGGTGTACACATGCATGGTGACCAATGTTGCAGGCAACTCCAACGCCTCG
 GCCTACCTCAATGTGAGCACGGCTGAGCTTAACACCTCCAACCTACAGCTTCTTCACCACAGT
 AACAGTGGAGACCACGGAGATCTCGCCTGAGGACACAACGCGAAAGTACAAGCCTGTTTCTTA
 CCACGTCCACTGGTTACCAGCCGGCATATACCACCTCTACCACGGTGCTCATTCAGACTACC
 CGTGTGCCCAAGCAGGTGGCAGTACCCGCGACAGACACCACTGACAAGATGCAGACCAGCCT
 GGATGAAGTCATGAAGACCACCAAGATCATCATTGGCTGCTTTGTGGCAGTGACTCTGCTAG
 CTGCCGCCATGTTGATTGTCTTCTATAAACTTCGTAAGCGGCACCAGCAGCGGAGTACAGTC
 ACAGCCGCCCGGACTGTTGAGATAATCCAGGTGGACGAAGACATCCCAGCAGCAACATCCGC
 AGCAGCAACAGCAGCTCCGTCCGGTGTATCAGGTGAGGGGGCAGTAGTGCTGCCACAAATTC
 ATGACCATATTAACCTACAACACCTACAAACCAGCACATGGGGCCCCACTGGACAGAAAACAGC
 CTGGGGAACCTCTCTGCACCCACAGTCACCACTATCTCTGAACCTTATATAATTACAGCCCA
 TACCAAGGACAAGGTACAGGAACTCAAATA**TGACT**CCCCCTCCCCCAAAAACTTATAAAAT
 GCAATAGAATGCACACAAAGACAGCAACTTTTGTACAGAGTGGGGAGAGACTTTTTCTTGTA
 TATGCTTATATATTAAGTCTATGGGCTGGTTAAAAAAACAGATTATATTAATAATTAAAGA
 CAAAAGTCAAAACA

FIGURE 157

MKLLWQVTVHHHTWNAILLPFVYLTAQVWILCAAIAAAASAGPQNCPSVCSCSNQFSKVVCT
RRGLSEVPQGIPSNTRYLNLMMENNIQMIQADTFRHLHHLEVLQLGRNSIRQIEVGAFNGLAS
LNTLELFDNWLTVIPSGAFEYLSKLRELWLRNNPIESIPSYAFNRVPSLMRLDLGELKKLEY
ISEGAFEGFLNLKYLNLGMCNIKDMPNLTPLVGLEELEMMSGNHFPETIRPGSFHGLSSLKKLW
VMNSQVSLIERNAFDGLASLVELNLAHNNLSSLPHDLFTPLRYLVELHLHHNPWNCDCDILW
LAWWLREYIPTNSTCCGRCHAPMHMRGRYLVEVDQASFQCSAPFIMDAPRDLNISEGRMAEL
KCRTPPMSSVKWLLPNGTVLSHASRHPRISVLNDGTLNFSHVLLSDTGVYTCMVTNVAGNSN
ASAYLNVSTAELNTSNYSFFTTVTVETTEISPEDTTRKYKPVPTTSTGYQPAYTTSTTVLIQ
TTRVPKQVAVPATDTTDMQOTSLDEVMTTKIIIGCFVAVTLLAAAMLIVFYKLRKRHQORS
TVTAARTVEIIQVDEDIPAATSAAATAAPSGVSGEGAVVLPTIHDHINYNTYKPAHGAHWTE
NSLGNSLHPTVTTISEPYIIQTHTKDKVQETQI

FIGURE 158

MELGCWTQLGLTFLQLLLISSLPREYTVINEACPGAENIMCRECEYDQIECVCPGKREVV
GYTIPCCRNEENECDSCLIHPGCTIFENCKSCRNGSWGGLTDDFYVKGIFYCAECRAGWYGGD
CMRCGQVLRAPKGQILLESYPLNAHCEWTIHAKPGFVIQLRFVMLSLEFDYMCQYDYVEVRD
GDNRDGQIIKRVCGNERPAPIQSIGSSLHVLFHSDGSKNFDGFGHAIYEEITACSSSPCFHDG
TCVLDKAGSYKCACLAGYTQRCENLLEERNCSDPGGPVNGYQKITGGPGLINGRHAKIGTV
VSFFCNNSYVLSGNEKRTCQQNGEWSGKQPICIKACREPKISDLVRRRVLPMPQVQSRETPLH
QLYSAAFSKQKLQSAPTKKPALPFGDLPMGYQHLHTQLQYECISPFYRRLGSSRRTCLRTGK
WSGRAPSCIPICGKIENITAPKTQGLRWPWQAAIYRRTSGVHDGSLHKGAWFLVCSGALVNE
RTVVVAAHCVTDLGKVTMIKTADLKVVLGKFYRDDDRDEKTIQSLQISAIILHPNYDPILLD
ADIAILKLLDKARISTRVQPICLAASRDLSTS FQESHITVAGWNVLADVRSPGFKNDTLRSG
VVSVDSSLCEEQHEDHGIPVSVTDNMFCASWEPTAPSDICTAETGGIAAVSFPGRASPEPR
WHLMGLVSWSYDKTCSHRLSTAFTKVL PFKDWIERNMK

FIGURE 160

ACCAGGCATTGTATCTTCAGTTGTCATCAAGTTCGCAATCAGATTGGAAAAGCTCAACTTGA
AGCTTTCTTGCCTGCAGTGAAGCAGAGAGATAGATATTATTACGTAATAAAAAAC**ATG**GGC
TTCAACCTGACTTTCCACCTTTCCTACAAATTCCGATTACTGTTGCTGTTGACTTTTGTGCCT
GACAGTGGTTGGGTGGGCCACCAGTAACTACTTCGTGGGTGCCATTCAAGAGATTCTTAAAG
CAAAGGAGTTCATGGCTAATTTCCATAAGACCCTCATTTTGGGGAAGGGAAAACTCTGACT
AATGAAGCATCCACGAAGAAGGTAGAAGTTGACAAGTGTCTTCTGTGTCTCCTTACCTCAG
AGGCCAGAGCAAGCTCATTTTCAAACCAGATCTCACTTTGGAAGAGGTACAGGCAGAAAATC
CCAAAGTGTCCAGAGGCCGGTATCGCCCTCAGGAATGTAAAGCTTTACAGAGGGTCGCCATC
CTCGTTCCCCACCGAACAGAGAGAAACACCTGATGTACCTGCTGGAACATCTGCATCCCTT
CCTGCAGAGGCAGCAGCTGGATTATGGCATCTACGTCATCCACCAGGCTGAAGGTAAAAAGT
TTAATCGAGCCAACTCTTGAATGTGGGCTATCTAGAAGCCCTCAAGGAAGAAAATTGGGAC
TGCTTTATATTCCACGATGTGGACCTGGTACCCGAGAATGACTTTAACCTTTACAAGTGTGA
GGAGCATCCCAAGCATCTGGTGGTTGGCAGGAACAGCACTGGGTACAGGTTACGTTACAGTG
GATATTTTGGGGGTGTTACTGCCCTAAGCAGAGAGCAGTTTTTCAAGGTGAATGGATTCTCT
AACAACTACTGGGGATGGGGAGGCGAAGACGATGACCTCAGACTCAGGGTTGAGCTCCAAAG
AATGAAAATTTCCCGGCCCTGCCTGAAGTGGGTAAATATACAATGGTCTTCCACACTAGAG
ACAAAGGCAATGAGGTGAACGCAGAACGGATGAAGCTCTTACACCAAGTGTCACGAGTCTGG
AGAACAGATGGGTGAGTAGTTGTTCTTATAAATTAGTATCTGTGGAACACAATCCTTTATA
TATCAACATCACAGTGGATTCTGGTTTTGGTGCA**TGA**CCCTGGATCTTTTGGTGATGTTTGG
AAGAACTGATTCTTTGTTTGCAATAATTTTGGCCTAGAGACTTCAAATAGTAGCACACATTA
AGAACCTGTTACAGCTCATTGTTGAGCTGAATTTTTCCTTTTTGTATTTTCTTAGCAGAGCT
CCTGGTGATGTAGAGTATAAAACAGTTGTAACAAGACAGCTTTCTTAGTCATTTTGTATCATG
AGGGTTAAATATTGTAATATGGATACTTGAAGGACTTTATATAAAAGGATGACTCAAAGGAT
AAAATGAACGCTATTTGAGGACTCTGGTTGAAGGAGATTTATTTAAATTTGAAGTAATATAT
TATGGGATAAAAGGCCACAGGAAATAAGACTGCTGAATGTCTGAGAGAACCAGAGTTGTTCT
CGTCCAAGGTAGAAAGGTACGAAGATACAATACTGTTATTATTTATCCTGTACAATCATCT
GTGAAGTGGTGGTGTGAGGTGAGAAGGCGTCCACAAAAGAGGGGAGAAAAGGCGACGAATCA
GGACACAGTGAAGTTGGGAATGAAGAGGTAGCAGGAGGGTGGAGTGTGGGCTGCAAAGGCAG
CAGTAGCTGAGCTGGTTGCAGGTGCTGATAGCCTTCAGGGGAGGACCTGCCCAGGTATGCCT
TCCAGTGATGCCACCAGAGAATACATTCTCTATTAGTTTTTAAAGAGTTTTTGTAAAATGA
TTTTGTACAAGTAGGATATGAATTAGCAGTTTACAAGTTTACATATTAATAATAATAATA
TGTCTATCAAATACCTCTGTAGTAAAATGTGAAAAAGCAAAA

FIGURE 161

MGFNLT FHLSYKFRL LLLLLL TLCLTVVGWATS NYFVGAIQEIPKAKEFMANFHKT LILGKGKT
LTNEASTKKVELDNCPSVSPYLRGQSKLIFKPDLTLEEVQAENPKVSRGRYRPQECKALQRV
AILVPHRNREKHLMYLLEHLHPFLQRQQLDYGIYVIHQAEKKFNRAKLLNVGYLEALKEEN
WDCFIFHDVDLVPENDFNLYKCEEHPKHLVVGRNSTGYRLRYSGYFGGVTALSREQFFKVNG
FSNNYWGWGGEDDDLRLRVELQRMKISRPLPEVGKYTMVFHTRDKGNEVNAERMKLLHQVSR
VWRDGLSSCSYKLVSV EHNPLYINITVDFWFGA

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites:

amino acids 4-7, 220-223 and 335-338

Xylose isomerase proteins:

amino acids 191-201

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 162

CGTGGGCGGGGTCGCGCAGCGGGCTGTGGGCGCGCCCGGAGGAGCGACCGCCGCAGTTCTC
 GAGCTCCAGCTGCATTCCCTCCGCGTCCGCCCCACGCTTCTCCCGCTCCGGGCCCCGCA**ATG**
 GCCCAGGCAGTGTGGTCGCGCCTCGGCCGCATCCTCTGGCTTGCCTGCCTCCTGCCCTGGGC
 CCCGGCAGGGGTGGCCGCAGGCCTGTATGAACTCAATCTCACCACCGATAGCCCTGCCACCA
 CGGGAGCGGTGGTGACCATCTCGGCCAGCCTGGTGGCCAAGGACAACGGCAGCCTGGCCCTG
 CCCGCTGACGCCACCTCTACCGCTTCCACTGGATCCACACCCCGCTGGTGCTTACTGGCAA
 GATGGAGAAGGGTCTCAGCTCCACCATCCGTGTGGTGGGCCACGTGCCCCGGGGAATTCCCGG
 TCTCTGTCTGGGTCACTGCCGCTGACTGCTGGATGTGCCAGCCTGTGGCCAGGGGCTTTGTG
 GTCCTCCCCATCACAGAGTTCCCTCGTGGGGGACCTTGTTGTCACCCAGAACACTTCCCTACC
 CTGGCCCAGCTCCTATCTCACTAAGACCGTCTGAAAGTCTCCTTCCCTCCTCCACGACCCGA
 GCAACTTCCCTCAAGACCGCCTTGTTTCTCTACAGCTGGGACTTCGGGGACGGGACCCAGATG
 GTGACTGAAGACTCCGTGGTCTATTATAACTATTCCATCATCGGGACCTTCACCGTGAAGCT
 CAAAGTGGTGGCGGAGTGGGAAGAGGTGGAGCCGGATGCCACGAGGGGCTGTGAAGCAGAAGA
 CCGGGGACTTCTCCGCTCGCTGAAGCTGCAGGAAACCTTCGAGGCATCCAAGTGTTGGGG
 CCCACCCTAATTACAGACCTTCCAAAAGATGACCGTGACCTTGAACCTTCTGGGGAGCCCTCC
 TCTGACTGTGTGCTGGCGTCTCAAGCCTGAGTGCCTCCCGCTGGAGGAAGGGGAGTGCCACC
 CTGTGTCCGTGGCCAGCACAGCGTACAACCTGACCCACACCTTCAGGGACCTTGGGGACTAC
 TGCTTCAGCATCCGGGCCGAGAATATCATCAGCAAGACACATCAGTACCACAAGATCCAGGT
 GTGGCCCTCCAGAATCCAGCCGGCTGTCTTTGCTTTCCCATGTGCTACACTTATCACTGTGA
 TGTTGGCCTTCATCATGTACATGACCTGCGGAATGCCACTCAGCAAAAGGACATGGTGGAG
 AACCCGGAGCCACCTCTGGGGTCAGGTGCTGCTGCCAGATGTGCTGTGGGCCTTTCTTGCT
 GGAGACTCCATCTGAGTACCTGGAATTTGTTGCTGAGAACCACGGGCTGCTCCCGCCCCCTCT
 ATAAGTCTGTCAAACTTACACCGT**GTAG**CACTCCCCCTCCCCACCCCATCTCAGTGTTAA
 CTGACTGCTGACTTGGAGTTTCCAGCAGGGTGGTGTGCACCACTGACCAGGAGGGGTTCAAT
 TGGCTGGGGCTGTTGGCCTGGATCATCCATCCATCTGTACAGTTCAGCCACTGCCACAAGCC
 CCTCCCTCTCTGTACCCCTGACCCAGCCATTCACCCATCTGTACAGTCCAGCCACTGACA
 TAAGCCCCACTCGGTTACCACCCCTTGACCCCTACCTTTGAAGAGGCTTCGTGCAGGACT
 TTGATGCTTGGGGTGTTCCGTGTTGACTCCTAGGTGGGCCTGGCTGCCCCACTGCCCATTCCT
 CTCATATTGGCACATCTGCTGTCCATTGGGGGTTCTCAGTTTCTCCCCCAGACAGCCCTAC
 CTGTGCCAGAGAGCTAGAAAGAAGGTCATAAAGGGTTAAAAATCCATAACTAAAGGTTGTAC
 ACATAGATGGGCACACTCACAGAGAGAAGTGTGCATGTACACACACCACACACACACACA
 CACACACACACAGAAATATAAACACATGCGTCACATGGGCATTTAGATGATCAGCTCTGTA
 TCTGGTTAAGTCGGTTGCTGGGATGCACCCTGCACTAGAGCTGAAAGGAAATTTGACCTCCA
 AGCAGCCCTGACAGGTTCTGGGCCCCGGGCCCTCCCTTTGTGCTTTGTCTCTGCAGTTCTTGC
 GCCCTTTATAAGGCCATCCTAGTCCCTGCTGGCTGGCAGGGGCCTGGATGGGGGGCAGGACT
 AATACTGAGTGATTGCAGAGTGCTTTATAAATATCACCTTATTTTATCGAAACCCATCTGTG
 AAACCTTTCAGTGAGGAAAAGGCCTTGACGCGGTAGAAGAGGTTGAGTCAAGGCCGGGCGCGG
 TGGCTCACGCCTGTAATCCAGCACTTTGGGAGGCCGAGGCGGGTGGATCACGAGATCAGGA
 GATCGAGACCACCTGGCTAACACGGTGAACCCCGTCTCTACTAAAAAATACAAAAAGTT
 AGCCGGGCGTGGTGGTGGGTGCCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATG
 GTGCGAACCCGGGAGGCGGAGCTTGCACTGAGCCCAGATGGCGCCACTGCACTCCAGCCTGA
 GTGACAGAGCGAGACTCTGTCTCCA

FIGURE 163

MAQAVWSRLGRILWLACLLPWAPAGVAAGLYELNLTTDSPATTGAVVTISASLVAKDNGSLA
LPADAHLYRFHWIHTPLVLTGKMEKGLSSTIRVVGHVPGEFPVSVWVTAADCWMCQPVARGF
VVLPITEFLVGDLVVTQNTSLPWSSYLTKTVLKVSFLLHDPNFKLTALFLYSWDFGDGTQ
MVTEDSVVYYNYSIIIGTFTVKLKVVAEWEEVEPDATRAVKQKTGDFSASLKLQETLRGIQVL
GPTLIQTFQKMTVTNLNFGSPPLTVCWRLKPECLPLEEGECHPVSVASTAYNLHTFRDPGD
YCFSIRAENIISKTHQYHKIQVWPSRIQPAVFAFPCATLITVMLAFIMYMTLRNATQQKDMV
ENPEPPSGVRCCCQMC CGPFLLLETPSEYLEIVRENHGLLPPLYKSVKTYTV

Important features of the protein:**Signal peptide:**

amino acids 1-24

Transmembrane domain:

amino acids 339-362

N-glycosylation sites.

amino acids 34-37, 58-61, 142-145, 197-200, 300-303 and 364-367

FIGURE 164

GCTCAAGACCCAGCAGTGGGACAGCCAGACAGACGGCACGATGGCACTGAGCTCCCAGATCT
GGGCCGCTTGCCCTCCTGCTCCTCCTCCTCGCCAGCCTGACCAGTGGCTCTGTTTTCCCA
CAACAGACGGGACAACCTTGCAGAGCTGCAACCCCAGGACAGAGCTGGAGCCAGGGCCAGCTG
GATGCCCATGTTCCAGAGGCGAAGGAGGCGAGACACCCACTTCCCCATCTGCATTTTCTGCT
GCGGCTGCTGTCATCGATCAAAGTGTGGGATGTGCTGCAAGACGTAGAACCTACCTGCCCTG
CCCCGTCCCCTCCCTTCCTTATTTATTCCTGCTGCCCCAGAACATAGGTCTTGGAATAAAA
TGGCTGGTTCTTTTGTTCACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAA

164/330

165/330

FIGURE 165

MALSSQIWAACLLLLLLLLASLTSGSVFPQQTGQLAELQPQDRAGARASWMPMFQRRRRRDTH
FPICIFCCGCCHRSKCGMCCKT

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 167

MFLLLPFDSLIVNLLGISLTVLFTLLLVFIIIVPAIFGVSFGIRKLYMKSLLKIFAWATLRME
RGAKEKNHQLYKPYTNGIIAKDPTSLEEEIKEIRRS GSSKALDNTPEFELSDIFYFCRKGME
TIMDDEVTKRFSAEELSWNLLSRTNYNFQYISLRLTVLWGLGVLIRYCFLPLRIALAF TG
ISLLVVGTTVVGYPNGRFKEFMSKHVHLMCYRICVRALTAIITYHDRENRPNGGICVANH
TSPIDV IILASDGY YAMVGQVHGGLMGVIQ RAMVKACPHVWFERSEVKDRHLVAKRLTEHVQ
DKSKLPILIFPEGTCINNTSVMMFKKGSFEIGATVYPVAIKYDPQFGDAFWNSSKYGMVTYL
LRMMTSWAIVCSVWYLP PMTREADED AVQFANRVKSAIARQGGLVDLLWDGGLKREKVKDTF
KEEQQKLYSKMIVGNHKDRSRS

167/330

FIGURE 168

GCCCCCTCGAAACCAGGACTCCAGCACCTCTGGTCCCGCCCTCACCCGGACCCCTGGGCCCTCA
CGTCTCCTCCAGGG**ATG**GCGCTGGCGGCTTTGATGATCGCCCTCGGCAGCCTCGGCCTCCAC
ACCTGGCAGGCCCAGGCTGTTCCACCATCCTGCCCCCTGGGCCTGGCTCCAGACACCTTTGA
CGATACCTATGTGGGTTGTGCAGAGGAGATGGAGGAGAAGGCAGCCCCCTGCTAAAGGAGG
AAATGGCCCACCATGCCCTGCTGCGGGAATCCTGGGAGGCAGCCAGGAGACCTGGGAGGAC
AAGCGTCGAGGGCTTACCTTGCCCCCTGGCTTCAAAGCCCAGAATGGAATAGCCATTATGGT
CTACACCAACTCATCGAACACCTTGTACTGGGAGTTGAATCAGGCCGTGCGGACGGGCGGAG
GCTCCCGGGAGCTCTACATGAGGCACTTTCCTTCAAGGCCCTGCATTTCTACCTGATCCGG
GCCCTGCAGCTGCTGCGAGGCAGTGGGGGCTGCAGCAGGGGACCTGGGGAGGTGGTGTTCGG
AGGTGTGGGCAGCCTTCGCTTTGAACCCAAGAGGCTGGGGGACTCTGTCCGCTTGGGCCAGT
TTGCCTCCAGCTCCCTGGATAAGGCAGTGGCCACAGATTTGGGGAGAAGAGGCGGGGCTGT
GTGTCTGCGCCAGGGGTGCAGCTAGGGTCACAATCTGAGGGGGCCTCCTCTCTGCCCCCCTG
GAAGACTCTGCTCTTGCCCCCTGGAGAGTTCCAGCTCTCAGGGGTTGGGCCCT**TGA**AAGTCCA
ACATCTGCCACTTAGGAGCCCTGGGAACGGGTGACCTTCATATGACGAAGAGGCACCTCCAG
CAGCCTTGAGAAGCAAGAACATGGTTCCGGACCCAGCCCTAGCAGCCTTCTCCCCAACCAGG
ATGTTGGCCTGGGGAGGCCACAGCAGGGCTGAGGGAACCTCTGCTATGTGATGGGGACTTCCT
GGGACAAGCAAGGAAAGTACTGAGGCAGCCACTTGATTGAACGGTGTTGCAATGTGGAGACA
TGGAGTTTTATTGAGGTAGCTACGTGATTAATGGTATTGCAGTGTGGA

FIGURE 169

MALAALMIALGSLGLHTWQAQAVPTILPLGLAPDTFDDTYVGCAEEMEEKAAPLLKEEMAHH
ALLRESWEAAQETWEDKRRGLTLPPGFKAQNGIAIMVYTNSSNTLYWELNQAVRTGGGSREL
YMRHFPPKALHFYLIRALQLLRGSGGCSRGPGEVVFRGVGSLRFEPKRLGDSVRLGQFASSS
LDKAVAHRFGEKRRGCVSAPGVQLGSQSEGASSLPPWKTL LLAPGEFQLSGVGP

169/330

FIGURE 170

GTGGCTTCATTTTCAGTGGCTGACTTCCAGAGAGCAATATGGCTGGTTCCCCAACATGCCTCA
CCCTCATCTATATCCTTTGGCAGCTCACAGGGTCAGCAGCCTCTGGACCCGTGAAAGAGCTG
GTCGGTTCCGTTGGTGGGGCCGTGACTTTCCCCCTGAAGTCCAAAGTAAAGCAAGTTGACTC
TATTGTCTGGACCTTCAACACAACCCCTCTTGTACCATAACAGCCAGAAGGGGGCACTATCA
TAGTGACCCAAAATCGTAATAGGGAGAGAGTAGACTTCCCAGATGGAGGCTACTCCCTGAAG
CTCAGCAAACCTGAAGAAGAATGACTCAGGGATCTACTATGTGGGGATATACAGCTCATCACT
CCAGCAGCCCTCCACCCAGGAGTACGTGCTGCATGTCTACGAGCACCTGTCAAAGCCTAAAG
TCACCATGGGTCTGCAGAGCAATAAGAATGGCACCTGTGTGACCAATCTGACATGCTGCATG
GAACATGGGGAAGAGGATGTGATTTATACCTGGAAGGCCCTGGGGCAAGCAGCCAATGAGTC
CCATAATGGGTCCATCCTCCCCATCTCCTGGAGATGGGGAGAAAGTGATATGACCTTCATCT
GCGTTGCCAGGAACCCTGTGAGCAGAACTTCTCAAGCCCCATCCTTGCCAGGAAGCTCTGT
GAAGGTGCTGCTGATGACCCAGATTCTCCATGGTCTCCTGTGTCTCCTGTTGGTGCCCCT
CCTGCTCAGTCTCTTTGTACTGGGGCTATTTCTTTGGTTTCTGAAGAGAGAGAGACAAGAAG
AGTACATTGAAGAGAAGAAGAGAGTGGACATTTGTGCGGGAACTCCTAACATATGCCCCCAT
TCTGGAGAGAACACAGAGTACGACACAATCCCTCACACTAATAGAACAATCCTAAAGGAAGA
TCCAGCAAATACGGTTTACTCCACTGTGGAAATACCGAAAAAGATGGAAAATCCCCACTCAC
TGCTCACGATGCCAGACACACCAAGGCTATTTGCCTATGAGAATGTTATCTAGACAGCAGTG
CACTCCCCTAAGTCTCTGCTCA

171/330

FIGURE 171

MAGSPTCLTLIYILWQLTGSAASGPVKELVGSVGGAVTFPLKSKVKQVDSIVWTFNTTPLVT
IQPEGGTIIIVTQNRNRERERVDFFPDGGYSLKLSKLKKNDSGIYYVGIYSSSLQQPSTQEYVLHV
YEHLSPKPKVTMGLQSNKNGTCVTNLTCCMEHGEEDVIYTWKALGQAANESHNGSILPISWRW
GESDMTFICVARNPVSRNFSSPILARKLCEGAADDPDSSMVLLCLLLVPLLLSLFLVLGLFLW
FLKRERQEYIEEKQRVDICRETPNICPHSGENTYDTPHTNRTILKEDPANTVYSTVEIP
KKMENPHSLLTMPDTPRLFAYENVI

171/330

FIGURE 172

CTGGTTCCCCAACATGCCTCACCCTCATCTATATCCTTTGGCAGCTCACAGGGTCAGCAGCC
TCTGGACCCCGTGAAAGAGCTGGTCGGTTCCGTTGGTGGGGCCGTGACTTTCCCCCTGAAGTC
CAAAGTAAAGCAAGTTGACTCTATTGTCTGGACCTTCAACACAACCCCTCTTGTCACCATAC
AGCCAGAAGGGGGCCTATCATAGTGACCCAAAATCGTAATAGGGAGAGAGTAGACTTCCCA
GATGGAGGCTACTCCCTGAAGCTCAGCAAATGAAGAAGAATGACTCAGGGATCTACTATGT
GGGGATATACAGCTCATCACTCCAGCAGCCCTCCACCCAGGAGTACGTGCTGCATGTCTACG
AGCACCTGTCAAAGCCTAAAGTCACCATGGGTCTGCAGAGCAATAAGAATGGCACCTGTGTG
ACCAATCTGACATGCTGCATGGAACATGGGGAAGAGGATGTGATTTATACCTGGAAGGCCCT
GGGGCAAGCAGCCAATGAGTCCCATAATGGGTCCATCCTCCCCATCTCCTGGAGATGGGGAG
AAAGTGATATGACCTTCATCTGCGTTGCCAGGAACCCTGTCAGCAGAACTTCTCAAGCCCC
ATCCTTGCCAGGAAGCTCTGTGAAGGTGCTGCTGATGACCCAGATTCCCTCCATGGTCCTCCT
GTGTCTCCTGTTGGTGCCCCCTCCTGCTCAGTCTCTTTGTACTGGGGCTATTTCTTTGGTTTC
TGAAGAGAGAGAGACAAGAAGAGTACATTGAAGAGAAGAAGAGAGTGGACATTTGTCGGGAA
ACTCCTAACATATGCCCCCATTTCTGGAGAGAACACAGAGTACGACACAATCCCTCACACTAA
TAGAACAATCCTAAAGGAAGATCCAGCAAATACGGTTTACTCCACTGTGGAAATACCGAAAA
AGATGGAAAATCCCCACTCACTGCTCACGATGCCAGACACACCAAGGCTATTTGCCTATGAG
AATGTTATCTAGACAGCAGTGCACTCCCCTAAGTCTCTGCTCAAAAAAAAAAAAAAAAAAAAA

FIGURE 173

GAAAGACGTGGTCCTGACAGACAGACAATCCTATTCCCTACCAAA**ATG**AAGATGCTGCTGCT
GCTGTGTTTGGGACTGACCCTAGTCTGTGTCCATGCAGAAGAAGCTAGTTCTACGGGAAGGA
ACTTTAATGTAGAAAAGATTAATGGGGAATGGCATACTATTATCCTGGCCTCTGACAAAAGA
GAAAAGATAGAAGAACATGGCAACTTTAGACTTTTTCTGGAGCAAATCCATGTCTTGGAGAA
TTCCTTAGTTCTTAAAGTCCATACTGTAAGAGATGAAGAGTGCTCCGAATTATCTATGGTTG
CTGACAAAACAGAAAAGGCTGGTGAATATTCTGTGACGTATGATGGATTCAATACATTTACT
ATACCTAAGACAGACTATGATAACTTTCTTATGGCTCACCTCATTAACGAAAAGGATGGGGA
AACCTTCCAGCTGATGGGGCTCTATGGCCGAGAACCAGATTTGAGTTCAGACATCAAGGAAA
GGTTTGCACAACTATGTGAGGAGCATGGAATCCTTAGAGAAAATATCATTGACCTATCCAAT
GCCAATCGCTGCCTCCAGGCCCGAGAATGAAGAATGGCCTGAGCCTCCAGTGTTGAGTGGAC
ACTTCTCACCAGGACTCCACCATCATCCCTTCCTATCCATACAGCATCCCCAGTATAAATTC
TGTGATCTGCATTCCATCCTGTCTCACTGAGAAGTCCAATTCCAGTCTATCAACATGTTACC
TAGGATACCTCATCAAGAATCAAAGACTTCTTTAAATTTCTCTTTGATACACCCTTGACAAT
TTTTCATGAAATTATTCCTCTTCCTGTTCAATAAATGATTACCCTTGCACTTAA

174/330

FIGURE 174

MKMLLLLCGLTLVCVHAEAEASSTGRNFNVEKINGEWHTIILASDKREKIEEHGNFRLFLEQ
IHVLENSLVLKVHTVRDEECSELSMVADKTEKAGEYSVTYDGENTFTIPKTDYDNFLMAHLI
NEKDGETFQIMGLYGREPDLSSDIKERFAQLCEEHGILRENIIDLSNANRCLQARE

174/330

FIGURE 175

GGCTCGAGCGTTTCTGAGCCAGGGGTGACCA**ATG**ACCTGCTGCGAAGGATGGACATCCTGCAA
TGGATTCAGCCTGCTGGTTCTACTGCTGTTAGGAGTAGTTCTCAATGCGATACCTCTAATTG
TCAGCTTAGTTGAGGAAGACCAATTTTCTCAAACCCCATCTCTTGCTTTGAGTGGTGGTTC
CCAGGAATTATAGGAGCAGGTCTGATGGCCATTCCAGCAACAACAATGTCCTTGACAGCAAG
AAAAAGAGCGTGCTGCAACAACAGAAGTGAATGTTTCTTTCATCATTTTTTCAGTGTGATCA
CAGTCATTGGTGCTCTGTATTGCATGCTGATATCCATCCAGGCTCTCTTAAAAGGTCCTCTC
ATGTGTAATTCTCCAAGCAACAGTAATGCCAATTGTGAATTTTCATTGAAAAACATCAGTGA
CATTCATCCAGAATCCTTCAACTTGCAGTGGTTTTTCAATGACTCTTGTGCACCTCCTACTG
GTTTCAATAAACCCACCAGTAACGACACCATGGCGAGTGGCTGGAGAGCATCTAGTTTCCAC
TTCGATTCTGAAGAAAACAAACATAGGCTTATCCACTTCTCAGTATTTTTAGGTCTATTGCT
TGTTGGAATTCTGGAGGTCCTGTTTGGGCTCAGTCAGATAGTCATCGGTTTCCTTGGCTGTC
TGTGTGGAGTCTCTAAGCGAAGAAGTCAAATTGTG**TAG**TTTAATGGGAATAAAATGTAAGTA
TCAGTAGTTTGAAAAAAAAAAAA

175/330

FIGURE 176

MTCCEGWTSCNGFSLLVLLLLGVVLNAIPLIVSLVEEDQFSQNPISCFEWWFPGIIGAGLMA
IPATTMSLTARKKRACNNRTGMFLSSFFSVITVIGALYCMLISIQALLKGPLMCNSPSNSNA
NCEFSLKNISDIHPESFNLQWFFNDSCAPPTGFNKPTSNDTMASGWRASSFHFDSEENKHRL
IHFSVFLGLLLLVGILEVLFGLSQIVIGFLGCLCGVSKRRSQIV

176/330

FIGURE 177

GTCGAATCCAAATCACTCATTGTGAAAGCTGAGCTCACAGCCGAATAAGCCACC**ATG**AGGCT
GTCAGTGTGTCTCCTGATGGTCTCGCTGGCCCTTTGCTGCTACCAGGCCCATGCTCTTGTCT
GCCCAGCTGTTGCTTCTGAGATCACAGTCTTCTTATTCTTAAGTGACGCTGCGGTAAACCTC
CAAGTTGCCAAACTTAATCCACCTCCAGAAGCTCTTGCAGCCAAGTTGGAAGTGAAGCACTG
CACCGATCAGATATCTTTTAAGAAACGACTCTCATTGAAAAAGTCCTGGTGGAAA**TAG**TGAA
AAAATGTGGTGTGTGACATGTAAAAATGCTCAACCTGGTTTCCAAAGTCTTCAACGACACC
CTGATCTTCACTAAAAATTGTAAAGGTTTCAACACGTTGCTTTAATAAATCACTTGCCCTGC

177/330

178/330

FIGURE 178

MRLSVCLLMVSLALCCYQAHALVCPAVASEITVFLFLSDAAVNLQVAKLNPPPEALAAKLEV
KHCTDQISFKKRLSLKKSWWK

178/330

FIGURE 179

ATCCGTTCTCTGCGCTGCCAGCTCAGGTGAGCCCTCGCCAAGGTGACCTCGCAGGACACTGG
TGAAGGAGCAGTGAGGAACCTGCAGAGTCACACAGTTGCTGACCAATTGAGCTGTGAGCCTG
GAGCAGATCCGTGGGCTGCAGACCCCCGCCCCAGTGCCTCTCCCCCTGCAGCCCTGCCCCCTC
GAACTGTGACATGGAGAGAGTGACCCTGGCCCTTCTCCTACTGGCAGGCCTGACTGCCTTGG
AAGCCAATGACCCATTTGCCAATAAAGACGATCCCTTCTACTATGACTGGAAAAACCTGCAG
CTGAGCGGACTGATCTGCGGAGGGCTCCTGGCCATTGCTGGGATCGCGGCAGTTCTGAGTGG
CAAATGCAAATACAAGAGCAGCCAGAAGCAGCACAGTCCTGTACCTGAGAAGGCCATCCCAC
TCATCACTCCAGGCTCTGCCACTACTTGCTTGAGCACAGGACTGGCCTCCAGGGATGGCCTGA
AGCCTAACACTGGCCCCCAGCACCTCCTCCCCCTGGGAGGCCTTATCCTCAAGGAAGGACTTC
TCTCCAAGGGCAGGCTGTTAGGCCCTTTCTGATCAGGAGGCTTCTTTATGAATTAAACTCG
CCCCACCACCCCCTCA

180/330

FIGURE 180

MERVTLALLLLAGLTALEANDPFANKDDPFYYDWKNLQLSGLICGGLLAIAGIAAVLSGKCK
YKSSQKQHSPVPEKAIP LITPGSATTC

180/330

FIGURE 181

GGAGAAGAGGTTGTGTGGGACAAGCTGCTCCCGACAGAAGGATGTCGCTGCTGAGCCTGCCC
TGGCTGGGCCTCAGACCGGTGGCAATGTCCCCATGGCTACTCCTGCTGCTGGTGTGGGCTC
CTGGCTACTCGCCCGCATCCTGGCTTGGACCTATGCCTTCTATAACAACCTGCCGCCGGCTCC
AGTGTTTCCCACAGCCCCCAAACGGAACCTGGTTTTGGGGTCACCTGGGCCTGATCACTCCT
ACAGAGGAGGGCTTGAAGGACTCGACCCAGATGTGGCCACCTATTCCCAGGGCTTTACGGT
ATGGCTGGGTCCCATCATCCCCTTCATCGTTTTATGCCACCCTGACACCATCCGGTCTATCA
CCAATGCCTCAGCTGCCATTGCACCCAAGGATAATCTCTTCATCAGGTTCTGAAGCCCTGG
CTGGGAGAAGGGATACTGCTGAGTGGCGGTGACAAGTGGAGCCGCCACCGTCGGATGCTGAC
GCCCGCCTTCCATTTCAACATCCTGAAGTCCTATATAACGATCTTCAACAAGAGTGCAAACA
TCATGCTTGACAAGTGGCAGCACCTGGCCTCAGAGGGCAGCAGTCGTCTGGACATGTTTGAG
CACATCAGCCTCATGACCTTGGACAGTCTACAGAAATGCATCTTCAGCTTTGACAGCCATTG
TCAGGAGAGGCCCAGTGAATATATTGCCACCATCTTGGAGCTCAGTGCCCTTGTAGAGAAAA
GAAGCCAGCATATCCTCCAGCACATGGACTTTCTGTATTACCTCTCCCATGACGGGCGGCGC
TTCCACAGGGCCTGCCGCCTGGTGCATGACTTCACAGACGCTGTCATCCGGGAGCGGCGTCG
CACCCTCCCCACTCAGGGTATTGATGATTTTTTTCAAAGACAAAGCCAAGTCCAAGACTTTGG
ATTTCAATTGATGTGCTTCTGCTGAGCAAGGATGAAGATGGGAAGGCATTGTGAGATGAGGAT
ATAAGAGCAGAGGCTGACACCTTCATGTTTGGAGGCCATGACACCACGGCCAGTGGCCTCTC
CTGGGTCTGTACAACCTTGCGAGGCACCCAGAATAACCAGGAGCGCTGCCGACAGGAGGTGC
AAGAGCTTCTGAAGGACCGGATCCTAAAGAGATTGAATGGGACGACCTGGCCCAGCTGCCC
TTCCTGACCATGTGCGTGAAGGAGAGCCTGAGGTTACATCCCCCAGCTCCCTTCATCTCCCG
ATGCTGCACCCAGGACATTGTTCTCCAGATGGCCGAGTCATCCCCAAAGGCATTACCTGCC
TCATCGATATTATAGGGGTCCATCACAACCCAACTGTGTGGCCGGATCCTGAGGTCTACGAC
CCCTTCCGCTTTGACCCAGAGAACAGCAAGGGGAGGTCACCTCTGGCTTTTATTCCTTTCTC
CGCAGGGCCCAGGAACTGCATCGGGCAGGCGTTCGCCATGGCGGAGATGAAAGTGGTCCTGG
CGTTGATGCTGCTGCACTTCCGGTTCCTGCCAGACCACACTGAGCCCCGCAGGAAGCTGGAA
TTGATCATGCGCGCCGAGGGCGGGCTTTGGCTGCGGGTGGAGCCCCTGAATGTAGGCTTGCA
GTGACTTTCTGACCCATCCACCTGTTTTTTTGCAGATTGTCATGAATAAAACGGTGCTGTCAAA

FIGURE 182

MSLLSLPWLGLRPVAMSPWLLLLLVVGSWLLARILAWTYAFYNNCRRLQCFPQPPKRNWFWG
HLGLITPTEEGLKDSTQMSATYSQGFTVWLGPIIPFIVLCHPDTIRSITNASAAIAPKDNLF
IRFLKPWLGE GILLSGGDKWSRHRRLTPAFHFNILKSYITIFNKSANIMLDKWQHLLASEGS
SRLDMFEHISLMTLDSLQKCIFSFDSHCQERPSEYIATILELSALVEKRSQHILQHMDFLYY
LSHDGRRFHRACRLVHDFTDAVIRERRRTLPTQGIDDFKDKAKSKTLDFIDVLLLSKDEG
KALSDEDIRAEADTFMFGGHDTTASGLSWVLYNLARHPEYQERCRQEVQELLKDRDPKEIEW
DDLAQLPFLTMCVKESLRLHPPAPFISRCCTQDIVLPDGRVIPKGITCLIDIIGVHHNPTVW
PDPEVYDPFRFDPENSKGRSPLAFIPFSAGPRNCIGQAFAMAEMKVVLALMLLHFRFLPDHT
EPRRKLELIMRAEGGLWLRVEPLNVGLQ

182/330

FIGURE 183

CAACAGAAGCCAAGAAGGAAGCCGTCTATCTTGTGGCGATCATGTATAAGCTGGCCTCCTGC
TGTTTGCTTTTCACAGGATTCTTAAATCCTCTCTTATCTCTTCCTCTCCTTGACTCCAGGGA
AATATCCTTTCAACTCTCAGCACCTCATGAAGACGCGCGCTTAACTCCGGAGGAGCTAGAAA
GAGCTTCCCTTCTACAGATATTGCCAGAGATGCTGGGTGCAGAAAGAGGGGATATTCTCAGG
AAAGCAGACTCAAGTACCAACATTTTTTAACCCAAGAGGAAATTTGAGAAAGTTTCAGGATTT
CTCTGGACAAGATCCTAACATTTTACTGAGTCATCTTTTGGCCAGAATCTGGAAACCATAACA
AGAAACGTGAGACTCCTGATTGCTTCTGGAAATACTGTGTCTTGAAGTGAAATAAGCATCTGT
TAGTCAGCTCAGAAACACCCATCTTAGAATATGAAAAATAACACAATGCTTGATTTGAAAAC
AGTGTGGAGAAAACTAGGCAAACTACACCCTGTTTATTGTTACCTGGAAAATAAATCCTCT
ATGTTTTGCACAAAAAAAAAAAAAAAAA

183/330

184/330

FIGURE 184

MYKLASCCLLFTGFLNPLLSLPLLDREISFQLSAPHEDARLTPEELERASLLQILPEMLGA
ERGDILRKADSSTNIFNPRGNLRKFQDFSGQDPNILLSHLLARIWKPYKKRETPDCFWKYCV

184/330

FIGURE 185

GAACATTTTTAGTTCCCAAGGAATGTACATCAGCCCCACGGAAGCTAGGCCACCTCTGGGAT
GGGGTTGCTGGTTTAAAACAAACGCCAGTCATCCTATATAAGGACCTGACAGCCACCAGGCA
CCACCTCCGCCAGGAAGTGCAGGCCCCACCTGTCTGCAACCCAGCTGAGGCC**ATG**CCCTCCCC
AGGGACCGTCTGCAGCCTCCTGCTCCTCGGCATGCTCTGGCTGGACTTGGCCATGGCAGGCT
CCAGCTTCCTGAGCCCTGAACACCAGAGAGTCCAGCAGAGAAAGGAGTCGAAGAAGCCACCA
GCCAAGCTGCAGCCCCGAGCTCTAGCAGGCTGGCTCCGCCCCGGAAGATGGAGGTCAAGCAGA
AGGGGCAGAGGATGAACTGGAAGTCCGGTTCAACGCCCCCTTTGATGTTGGAATCAAGCTGT
CAGGGGTTTCAGTACCAGCAGCACAGCCAGGCCCTGGGGAAGTTTCTTCAGGACATCCTCTGG
GAAGAGGCCAAAGAGGCCCCAGCCGACAAG**TGA**TCGCCCACAAGCCTTACTCACCTCTCTCT
AAGTTTAGAAGCGCTCATCTGGCTTTTCGCTTGCTTCTGCAGCAACTCCCACGACTGTTGTA
CAAGCTCAGGAGGCGAATAAATGTTCAAACCTGTA

186/330

FIGURE 186

MPSPGTVCLLLLGMLWLDLAMAGSSFLSPEHQRVQQRKESKKPPAKLQPRALAGWLRPEDG
GQAEGAEDELEVRFNAPFDVGIKLSGVQYQQHSQALGKFLQDILWEEAKEAPADKO

186/330

FIGURE 187

CGGCCACAGCTGGCATGCTCTGCCTGATCGCCATCCTGCTGTATGTCCTCGTCCAGTACCTC
GTGAACCCCGGGGTGCTCCGCACGGACCCAGATGTCAAGAAT**ATGA**ACACGTGGCTGCTGT
TCCTCCCCCTGTTCCCGGTGCAGGTGCAGACCCTGATAGTCGTGATCATCGGGATGCTCGTG
CTCCTGCTGGACTTTCTTGCTTGGTGCACCTGGGCCAGCTGCTCATCTTCCACATCTACCT
GAGTATGTCCCCCACCCTAAGCCCCCGATCCCCCAAGGCTGGGTGGTCAGAGCTGCTCATC
TTACACCTCTACTTGAGTATGTCCCTAACCTGAGCCCCCACGCCTGGGGCCAGAGTCTTT
GTCCCCCGTGTGCGCATGTGTTTCAAGGTCAGCCTCTCCAGAAGTGAGATCATGGACAAAAA
GGGCAAATCACAGGAAGAAATTAAATCCATGAGGACCCAGCAGGCCCCAGCAAGAAGCTGAAC
TCACGCCGAGACCTGCAGGAGTGGTGCCAGGTGCT**TGA**AGTAACAAGTTTAAAATGTTTCTCAGA
GACAATGGAATGGAATCTATTAGGCAAGAACAGGACATTATGAAATAAGGACAGGTGGACTT
CCAAAAACACAAGTAGAAATTCTAACAATGAAATATATTACAGGCAGGTACCCACTAACCA
AACAACTGAAGCGAGAGCTGTGGTCTTGCTTGGTCTCACAGTGGGCACAGCGGTAGGCGGTC
AGTCATGTTGCTGAACGACGGAGGGTAAACTCCCCAGCCCCAAGAAAACCTGTGTTGGAAGT
AACAAACAACCTCCCTGCTCCTGGCACCAGCCGTTTTGGTTCATGGTGGGCCAGCTGCAAAGCG
TCTTCCATTCTCTGGGCAGTGGTGGCCCCGAGGCTGTGGCCTCTCAGGGGGTTTCTGTGGAC
ACGGGCAGCAGAGTGTGTCCAGGCCAGCCCCAAGAATGCCCTGCTCCTGACAGCTTGGCCA
ACCCCTGGTCAGGGCAGAGGGAGTTGGGTGGGTGAGGCTCTGGGCTCACCTCCATCTCCAGA
GCATCCCCCTGCCTGCAGTTGTGGCAAGAACGCCAGCTCAGAATGAACACACCCCCACCAAGA
GCCTCCTTGTTTATAACCACAGGTACCCTACAAACCACTGTCCCCACACAACCTGGGGAT
GTTTTTAAACACACACCTCTAACGCATATCTTACAGTCACTGTTGTCTTGCCTGAGGGTTGA
ATTTTTTTTAAATGAAAGTGCAATGAAAATCACTGGATTAAATCCTACGGACACAGAGCTGAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 189

GGAGTGCAGATGGCATCCTTCGGTTCTTCCAGACAAGCTGCAAGACGCTGACC**ATG**GCCAAG
ATGGAGCTCTCGAAGGCCTTCTCTGGCCAGCGGACACTCCTATCTGCCATCCTCAGCATGCT
ATCACTCAGCTTCTCCACAACATCCCTGCTCAGCAACTACTGGTTTGTGGGCACACAGAAGG
TGCCCAAGCCCCTGTGCGAGAAAGGTCTGGCAGCCAAGTGCTTTGACATGCCAGTGTCCCTG
GATGGAGATACCAACACATCCACCCAGGAGGTGGTACAATACTGGGAGACTGGGGATGA
CCGGTTCTCCTTCCGGAGCTTCCGGAGTGGCATGTGGCTATCCTGTGAGGAACTGTGGAAG
AACCAGGGGAGAGGTGCCGAAGTTTCATTGAACTTACACCACCAGCCAAGAGAGGTGAGAAA
GGACTACTGGAATTTGCCACGTTGCAAGGCCCATGTCACCCCACTCTCCGATTTGGAGGGAA
GCGGTTGATGGAGAAGGCTTCCCTCCCCTCCCCTCCCTTGGGGCTTTGTGGCAAAAATCCTA
TGGTTATCCCTGGGAACGCAGATCACCTACATCGGACTTCAATTTCATCAGCTTCCTCCTGCT
ACTAACAGACTTGCTACTCACTGGGAACCTGCCTGTGGGCTCAAACCTGAGCGCCTTTGCTG
CTGTTTCCTCTGTCCTGTCAGGTCTCCTGGGGATGGTGGCCACATGATGTATTACAAAGTC
TTCCAAGCGACTGTCAACTTGGGTCCAGAAGACTGGAGACCACATGTTTGGAAATTATGGCTG
GGCCTTCTACATGGCCTGGCTCTCCTTCACCTGCTGCATGGCGTCGGCTGTCACCACCTTCA
ACACGTACACCAGGATGGTGCTGGAGTTCAAGTGCAAGCA**TAG**TAAGAGCTTCAAGGAAAAC
CCGAAGTGCCTACCACATCACCATCAGTGTTTCCCTCGGCGGCTGTCAAGTGCAGCCCCCAC
CGTGGGTCCTTTGACCAGCTACCACCAGTATCATAATCAGCCCATCCACTCTGTCTCTGAGG
GAGTCGACTTCTACTCCGAGCTGCGGAACAAGGGATTTCAAAGAGGGGCCAGCCAGGAGCTG
AAAGAAGCAGTTAGGTCATCTGTAGAGGAAGAGCAGTGTTAGGAGTTAAGCGGGTTTGGGGA
GTAGGCTTGAGCCCTACCTTACACGTCTGCTGATTATCAACATGTGCTTAAGCCAACATCCG
TCTCTTGAGCATGGTTTTTTAGAGGCTACGAATAAGGCTATGAATAAGGGTTATCTTTAAGTC
CTAAGGGATTCTGGGTGCCACTGCTCTCTTTTCTCTACAGCTCCATCTTGTTTACCCAC
CCCACATCTCACACATCCAGAATTCCTTCTTTACTGATAGTTTCTGTGCCAGGTTCTGGGC
TAAACCATGGAGATAAAAAGAAGAGTAAAATACACTTCCCGACCTTAAGGATCTGAAA

190/330

FIGURE 190

MAKMELSKAFSGQRTLLSAILSMLSLSFSTTSLLSNYWFVGTQKVPKPLCEKGLAAKCFDMP
VSLDGD TNTSTQE VVQYNWETGDDRF SFRSFRSGMWLSCEETVEEPGERCRSFIELTPPAKR
GEKGLLEFATLQGPCHPTLRFGGKRLMEKASLPSPPLGLCGKNPMVIPGNADHLHRTSIHQL
PPATNRLATHWEPCLWAQTERLCCCFLCPVRSPGDGGPHDVFTSLPSDCQLGSRRL ETTCLE
LWLGLLHGLALLHLLHGVGCHHLQHVHQDGAGVQVQA

190/330

FIGURE 191

AACTGGAAGGAAAGAAAGAAAGGTCAGCTTTGGCCCAG**ATG**TGGTTACCCCTTGGTCTCCTG
TCTTTATGTCTTTCTCCTCTTCCTATTCTGTCATCTCCCTCACTTAAGTCTCAGGCCTGTCA
GCAGCTCCTGTGGACATTGCCATCCCCCTCTGGTAGCCTTCAGAGCAAACAGGACAACCTATG
TTATGGATGTTTTCCACCAACCAGGGTAGTGGCATGGAGCACCGTAACCATCTGTGCTTCTGT
GATCTCTATGACAGAGCCACTTCTCCACCTCTGAAATGTTCCCTGCTCTGAAATCTGGCATG
AGATGGCACAGGTGACCACGCAGAAGCCACCAGAATCTTGCTGCCCCTATTCTCCTCCCAA
GTCTGTTCTCTTATTGTCAACCTCAGCACAAACAGGCTGGCGCCAATGGCATTACAGAGAAAG
CAATCTGTGTGGCTAGTGGGCAGATTACCATGCAAGCCCCAGGAGAAATGGAGGAGCTTTGT
AGCCACCTCCCTGTCAGCCAGTATTAACATGTCCCCCTTCCCCCTGCCCCGCCGTAGATTGAG
GACATTCGCCCCCTGTGTGCCACCAAACCAGGACTTTCCCCCTTGGCTTGGCATCCCTGGCTCT
CTCCTGGTACCCAGCAAGACGTCTGTTCCAGGGCAGTGTAGCATCTTTCAAGCTCCGTTACT
ATGGCGATGGCCATGATGTTACAATCCCACTTGCCTGAATAATCAAGTGGGAAGGGGAAGCA
GAGGGAAATGGGGCCATGTGAATGCAGCTGCTCTGTTCTCCCTACCCTGAGGAAAAACCAA
GGGAAGCAACAGGAACCTTCTGCAACTGGTTTTTATCGGAAAGATCATCCTGCCTGCAGATGC
TGTTGAAGGGGCACAAGAAATGTAGCTGGAGAAGATTGATGAAAGTGCAGGTGTGTAAGGAA
ATAGAACAGTCTGCTGGGAGTCAGACCTGGAATTCTGATTCCAACTCTTTATTACTTTGGG
AAGTCACTCAGCCTCCCCGTAGCCATCTCCAGGGTGACGGAACCCAGTGTATTACCTGCTGG
AACCAAGGAAACTAACAATGTAGGTTACTAGTGAATACCCCAATGGTTTCTCCAATTATGCC
CATGCCACCAAACAATAAAACAAAATTCTCTAACACTGAAA

192/330

FIGURE 192

MWLPLGLLSLCLSPLPILSSPSLKSQACQQLLWTLPSPLVAFRANRTTYVMDVSTNQSGME
HRNHLCFCDLYDRATSPPLKCSLL

192/330

FIGURE 193

GTAGCGCGTCTTGGGTCTCCCGGCTGCCGCTGCTGCCGCCGCCCTCGGGTCGTGGAGCCAGGAGCGACGTCA
 CCGCC**ATG**GCAGGCATCAAAGCTTTGATTAGTTTGTCTTTGGAGGAGCAATCGGACTGATGTTTTGATGCTT
 GGATGTGCCCTTCCAATATACACAAATACTGGCCCTCTTTGTTCTATTTTTTTTACATCCTTTCACCTATTCC
 ATACTGCATAGCAAGAAGATTAGTGGATGATACAGATGCTATGAGTAACGCTTGTAAGGAACCTTGCCATCTTTC
 TTACAACGGGCATTTGTCGTGTGAGCTTTTGGACTCCCTATTGTATTTGCCAGAGCACATCTGATTGAGTGGGGA
 GCTTGTGCACTTGTCTCACAGGAAACACAGTCATCTTGCAACTATACTAGGCTTTTTCTTGGTCTTTGGAAG
 CAATGACGACTTCAGCTGGCAGCAGTGG**TGA**AAAGAAATTACTGAACTATTGTCAAATGGACTTCCTGTCAATTT
 GTTGGCCATTCACGCACACAGGAGATGGGGCAGTTAATGCTGAATGGTATAGCAAGCCTCTTGGGGGTATTTTA
 GGTGCTCCCTTCTCACTTTTATTGTAAGCATACTATTTTTCACAGAGACTTGCTGAAGGATTAAAAGGATTTTCT
 CTTTTGGAAAAGCTTGACTGATTTTCACACTTATCTATAGTATGCTTTTTGTGGTGTCTGCTGAATTTAAATAT
 TTATGTGTTTTTCTGTTAGGTTGATTTTTTTTTTGAATCAATATGCAATGTTAAACACTTTTTTAATGTAATCA
 TTTGCATTGGTTAGGAATTCAGAATTCGCCCGGCTCTATTACTGGTCAAGTACATCTTTTCTCTTAAAATTATT
 TAGCCTCCATTATTACAAAAAATTATAAAAAATAAGTTTTTCAGTCAGTCAGGATGACATCACTCCCAATGTTATG
 CAGACATACAGACGGTTGGCATACTGTATAGACTGTATACTCAGTGCAAATATAGCTGCATTTATACCTCAGAG
 GGGCCAAGTGTTAATGCCCATGCCCTCCGTTAAGGGTTGTTGGTTTTACTGGTAGACAGATGTTTTGTGGATTG
 AAAATTATTTTATGGAATTGCTACAGAGGAGTGCTTTTCTTCAATTGTTAGAAGAATTTATGTTAAACTTTA
 AGGTAAGGGTGTA AAAACATTTTTGAGATAAGGTTTTTATTTATGTTTATTATTGTTAGAGTGAGTTGCAATGT
 GGGAAAGAAATGACATTGAAATTCCAGTTTTTGAATCCTGTTTCTATTTATAAGTGAAATTTGTGATCTCCTATC
 AACCTTTCATGTTTTTACCCTGTTAAAATGGACATACATGGAACCACTACTGATGAGGGACAGTTGTATGTTTGC
 ATCATATATGCCAGAAAACCTTCCCTCTGCTTCCCTCTTTGACTTATTTGGTATGTTGTATATATTACATAAAA
 TAACTTTTCAAATATAGTTTAATAACACTTAGAAGTGTTTACTTACCTGGAAAATAATTGCTATGCCGTACATT
 CAGAGTGCCCCCTCCCCTGCAAGGCCTTGCCATGATTAACAAGTAACCTGTTAGTCTTACAGATAATTCATGCA
 TTAACAGTTTAAGATTTAGACCATGGTAATAGTAGTTCCTTATTCTCTAAGGTTATATCATATGTAATTTAAAAG
 TATTTTAAAGACAAGTTTCTGTATACCTCTGAAGTGTGTTTGTGTTTGTGAGTTCATCATGATAGATCTGCTGTTT
 CCTTATAAAAGGCATTTGTTGTGTGAGTTAATGCAAAGTAGCCAAGTCCAGCTATATAGCAGCTTCAGAAACAT
 ACCTGACCAAAAAATTTCCAGTAACCAGGCATGATCAATTTATAGTGGTCGTTTACATCTAATAATTATCAGGA
 CTTTTTTCAGGAGTGGGTTATAAAAAACATTCAAGTTGGTCTGACAGTATTTTGTAAAGGATATTTGTTTGTATG
 TTTATTCAGTATACTTACATAAAAAATTATTTGCCATCAGCCAAAACCTCAGTAATCATGACAGCTGTCTGTTGT
 TTTATGAAGTTTATTTCTCAAGAAAATGGGAATAAATTTGGGATTTGTTTCAGCTTTTTTACTAAAGATGCCTAA
 AGCCACAGGTTTTATTGCTTAACCTTAAGCCATGACTTTTAGATATGAGATGACGGGAAGCAGGACGAAATATCG
 GCGTGTGGCTGGAGCCTTCCCACTGGAGGCTGAAAGTGGCTTGTGGTATTATAATGTTTCAGATTTCAAGAGGAA
 GGTGCAGGTACACATGAGTTAGAGAGCTGGTGAGACAGTTGGGAACCTTTTGTGCTTGTGATCTACTGGACTTT
 TTTTTTGCAAGGAAGTGCAATTTCTGGTCTTCCCTATTTTCTGTCTGGATGTCAGTGCAGTGCAGTGCAGTGCAGTGC
 TTTTATCCACTTGGCCACAGACTTTTTCTAACAGCTGCGTATTATTTCTATATACTAATTGCATTGGCAGCATT
 GTGCTTTGACCTTGTATACTAGCTTGACATAGTGTCTCTGATTTCTAGGCTAGTTACTTGAGATATGAAT
 TTTCCATAGAATATGCACTGATACAACATTACCATTCTTCTATGGAAAGAAAACCTTTTGATGATGAAACAATAA
 AGATTTTAAATATCTATTTTAAAAA

194/330

FIGURE 194

MAGIKALISLSFGGAIGLMFLMLGCALPIYNKYWPLEVLEFFYILSPIPYCIARRLVDDTDAM
SNACKELAIFLTGTGIVVSAFGLPIVFARAHLEWGACALVLTGNTVIFATILGFFLVFGSND
DFSWQQW

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

CCACGCGTCCGCCACGCGTCCGCCACGCGTCCGCCACGCGTCCGCCACGCGTCCGCCACGCGTCCGCC
CACGCGTCCGGTGCAGCTCGCGCCGCACTGCTGCTGGAGGAGAGGAGCCGCGCGCTCTCGCGCTCCC
CGCGCGCGGTCCGCACTCCCCACCGCCCGCCCGCCCGCCGCGCCGCAAAAGCATGAGTAGAGCCGCTC
TTCGAGCTGCCCGGGGCGGAATGGCAGGCTGTTTCGGCGGAGTAAAGGTGGCGCCGGTCAGTGGTCGTTTC
CAATGACGGACATTAACCAGACTGTTCAGATCCTGGGGAGTCGCGAGCCCCGAGTTTGGAGTTTTCCTCCCCAC
AAGCTCAGACTCCGAAGTCAGAGAGGAAAGGAGCGGAGGAGGCAAGCTCGGGCTCCGGCAGCTAGTTGG
GAAACTTGGCGGTCTTAGAAGTCGCCTCCCGCCTTGCCGCGCCCTTGACGCCCCGAGCCGAGCAGCAAAGT
GAGACATTGTGCGCCTGCCAGATCCGCGCGCGCGGACCGGGGCTGCCTCGGAAACACAGAGGGGTCTTCTCTC
GCCCTGCATATAATTAGCCTGCACAAAGGAGCAGCTGAATGGAGTTGTCACTCTCTGAAAGAGATTCTC
GACCGAGCGCTTCCAATTGGACATTCTCCAGTCTCTCTGAAAGATTCTCGTAATGATTCTCTGCTGCTCGGT
CTCTGTCTATACTGGCTGCTGAGGAGGCCCTCGGGGTGGTCTTGTGTCTGTCTGGGGCCTGCTTTCAGATGCT
GCCCGCGCGCCCGACGGGGTGCCGCGAGCTGTGCCGTCGAGGGGCGGCTGCTGCTACTGCGAGGCGCTCAAC
TCACCGAGGCGCCCCACAACCTGTCCGCGCTGCTGGGCTTGTCCCTGCGCTACAACAGCCTCTCGGAGCTGCGC
CGCGGCCAGTTCACGGGGTTAATGCAGCTCACGTGGCTCTATCTGGATACAATCACATCTGCTCCGTGCAGGG
GGACGCTTTTCAGAACTGCGCCGAGTTAAGGAATCACGTGAGTTCCAACCATCACCACTGCCAACTGCCAA
CCACCTTCGGGCCCATGCCAACTGCGCAGCTGGACCTCTCGTACAACAAGCTGCAGGCGCTCGCGCCGAC
CTCTTCACGGGCTGCGGAAGCTCACCACGCTGCATATGCGGGCCAACGCCATCCAGTTTGTGCCCGTGCGCAT
CTTCCAGGAGCTCGCGCAGCCTCAAGTTTCTGCATATCGGATACAATCAGCTCAAGAGTCTGGCGCGCACTCTT
TCGCGCGCTTGTTTAAGCTCACCGAGCTGCACCTCGAGCACAACGACTTGGTCAAGGTGAACCTTCGCCCACTTC
CGCGCGCTCATCTCCCTGCACTCGCTCTGCCTGCGGAGGAACAAGGTGGCCATTGTGGTCAGCTCGCTGGACTG
GGTTTGGAACTTGGAGAAATGGACTTGTGCGGCAACGAGATCGAGTACATGGAGCCCCATGTGTTTCGAGACCG
TGCGCGCACTGCAGTCCCTGCAGCTGGACTCAACCGCCTCACCTACATCGAGCCCGGATCTCAACTCTTGG
AAGTCCCTGACAAGCATCACCTTGGCGGGGAACCTGTGGGATTGCGGGCGCAACGTGTGTGCCCTAGCCTCGTG
GCTCAGCAACTTCAGGGGCGCTACGATGGCAACTTGCAGTGCGCCAGCCGGAGTACGCACAGGGCGAGGACG
TCCTGGAGCCGCTGTACGCTTCCACCTGTGCGAGGATGGGGCGAGCCACCGAGCGCCACCTGCTCTCGGC
TCTACCAACCGCAGTGATCTGGGGCCCCCTGCCAGCTCGGCCACCACGCTCGCGCAGCGCGGGGAGGGGACGA
CGAGCGCACATTGAGCCCTGCCACCGTGGCTCTTCAGGCGCGCAGCAGCCGAGAACCGCGCTGCATCCACA
AGGTGCTCAGGGCACCATTGGCCCTCATCTTCTCTCTCTCATCTGGTCTGTGCTCTACGTGCTCTGGAAG
TGTTTCCAGCCAGCCTCAGGAGCTCAGACAGTGTGTTGTACGCGAGCGCAGGAAGCAAAAGCAGAAACAGAC
CATGCATCAGATGGCTGCCATTGTCTGCCAGGAATACTACGTTGATTACAACCCGAACCATATTGAGGGAGCCC
TGGTGATCATCAACGAGTATGGCTCGTGTACCTGCCACCGAGCCCGCGAGGGAATGCGAGGTGTAATGTCC
CAGTGGCTCTCAACCCATGCGCTACCAATACGCTTGGGCGAGCCGGGACGGGCCGGCGGGCACCAGGCTGGGT
CTCCTTGTCTGTGCTCTGATATGCTCCTTGACTGAACTTTAAGGGGATCTCTCCAGAGACTTGACATTTTAG
CTTTATTGTGTCTTAAAAAAGCAAAAGCAATTAACACACAACCAAAACCCCAACCAACCTTCAGGACAGTC
TATCTTAAATTTTATATGAGAATCCTTCTCCCTTTGAAGATCTGTCCATATTAGGAATCTGAGAGTGTAAA
AAAGGTGGCCATAGACAGAGAGAGAATAATCGTGCTTTGTTTATGCTACTCTCCCAACCTTGGCCATGATTA
AACATCATGTATGTAGAAGATCTTAAGTCCATACGCAATTCATGAAGAACCATTGAAAGAGGAATCTGCAATC
TGGGAGCTTTAAGAGCAAAATGATGACCATAGAAAAGCTATGTTCTTACTTTGTGTGTGTGTCTGTATGTTCTCGG
TTGTGTGTCTTTGTAGGCAAGCAACGTTGTCTACACAAACGGGAATTTAGTCCATATTTATGCCCCCTGT
GCCTCTAGCTCTGGAGATTTGGTGGGGGAGTGGGGGGAACCGAGGAATAAGGGAAGTGTAGTTTAACT
AAGGTTTGTGAACACTGAAATCTTTTCTTTCTCAAATTAATTATCTTTAAGCTTCAAGAACTTGCTCTGACC
CCTCTAAGCAAACTACTAAGCATTTAAAGAGGAATCTAATTTTAAAGGTGTAGCACCCTTTTTTTTATCTTC
CCACAGAGGGTGCTAATCTATTATGCTATGCTATCTGAAAAGAATTAAGGCCACAATTCAGTCTCGTCTCG
GGCATTTGTATGATTGACCTCCATTTGCAGTACCTTCCAGCTGATTAAAGTTCAGCAGTGGTATTGAGGTT
TTTTGAATATTTATATAGAAAAAAGCTTTTTCATGACAAATGACACTCTCACACAGTCTTAGCCCTAGTA
GTTTTTTAGGTTGGACAGGGAAGCAGGTTAAATGAGACCTGTCTCTGTCTGCATCAGAAAAAATAGGCAGT
CCCTGATGCTCAGATCTTAGCCTTGATATTAATAGTTGAGACCACCTACCCACAATGCAGCCTATACTCCAAG
ACTACAAAGTTACCATCGAAAGGAAGGTTATTCAGTAAAGGAATAGTTTCTCAACCATTTAAAAATAT
TCTTCTGAACTCATCAAAGTAGAAGAGCCCCCAACCTTTTCTCTGCTTCAAGAAGCAGACATTTGGTATG
ATTTAGCATCAACAACATTTATGAGTATATGTAAGTAATCAGAGGGGCAAAATGCCACTTGTTATTCTCCCA
AGTTTTTCAAGCAAGTACACAGATCTCTGGTAGGATTAGGGGCCACTGTGTTTCCGGCTATTTTATGTCCGA
CTTGTGACGAAAGTTTGTAGCTCTATCTGACATGGCCAGTGAAGCAGGCAATGATGATGATGATGATGATGAT
GGTAGAAGGAACATCATCATATACCCCTCTCACAGAGAAAATTATCAAAGAACCAGAAATATATCTGTTTTGG
AGCAAGAGTGTATAATGTTTCAGGGTAGTCAAATAAACATAAATTATCTCTCTAGATGAGTGGCGATGTTG
GCTGATTGTTGGTCTGCCATTGACAGAATGCAAAATAAAAGGAATTAGCTAGAATATGACCAATTAATGTGCTT
CTGAAATATATTTTGAATAGGTTTGAATGTCA

FIGURE 196

MDFLLLGLCLYWLLRRPSGVVLCLLGACFQMLPAAPSGCPQLCRCEGRLLYCEALNLTEAPH
NLSGLLGLSLRYNSLSELRAGQFTGLMQLTWLYLDHNNHICSVQGDAFQKLRRVKELTLSSNQ
ITQLPNTTFRPMPNLRSDLSYNKLQALAPDLFHGLRKLTTLHMRANAIQFVPVRI FQDCRS
LKFLDIGYNQLKSLARNSFAGLFKLTELHLEHNDLVKVNFAHFPRILISLHSLCLRRNKVAIV
VSSLDWVWNLEKMDLSGNEIEYMEPHVFETVPHLQSLQLDSNRLTYIEPRILNSWKSLSIT
LAGNLWDCGRNVCALASWLSNFQGRYDGNLQCASPEYAQGEDVLDVYAFHLCEDGAEPSTG
HLLSAVTNRSDLGPPASSATTLADGGEGQHDGTFEPATVALPGGEHAENAVQIHKVVTGTMA
LIFSFLIVVLVLYVSWKCFPASLRQLRQCFVTQRRKQKQKQTMHQMAAMSAQEYYVDYKPNH
IEGALVIINEYGSC TCHQQPARECEV

196/330

FIGURE 197

GTGCAAGGAGCCGAGGCGAGATGGGCGTCCTGGGCCGGGTCTGCTGTGGCTGCAGCTCTGC
GCACTGACCCAGGCGGTCTCCAAACTCTGGGTCCCCAACACGGACTTCGACGTCGCAGCCAA
CTGGAGCCAGAACCGGACCCCGTGCGCCGGCGGCCGCGTTGAGTTCCCGGCGGACAAGATGG
TGTCAGTCCTGGTGCAAGAAGGTCACGCCGTCTCAGACATGCTCCTGCCGCTGGATGGGGAA
CTCGTCCTGGCTTCAGGAGCCGGATTTCGGCGTCTCAGACGTGGGCTCGCACCTGGACTGTGG
CGCGGGCGAACCTGCCGTCTTCCGCGACTCTGACCGCTTCTCCTGGCATGACCCGCACCTGT
GGCGCTCTGGGGACGAGGCACCTGGCCTCTTCTTCGTGGACGCCGAGCGCGTGCCCTGCCGC
CACGACGACGTCTTCTTTCCGCCTAGTGCCTCCTTCCGCGTGGGGCTCGGCCCTGGCGCTAG
CCCCGTGCGTGTCGCGAGCATCTCGGCTCTGGGCCGGACGTTACGCGCGACGAGGACCTGG
CTGTTTTCTGGCGTCCCGCGCGGGCCGCTACGCTTCCACGGGCCGGGCGCGCTTGAGCGTG
GGCCCCGAGGACTGCGCGGACCCGTCTGGGCTGCGTCTGCGGCAACGCGGAGGCGCAGCCGTG
GATCTGCGCGGCCCTGCTCCAGCCCCCT

198/330

FIGURE 198

MGVLGRVLLWLQLCALTQAVSKLWVPNTDFDVAANWSQNRTPCAGGAVEFPADKMVSVLVQE
GHAVSDMLLPLDGELVLASGAGFGVSDVGSHLDCGAGEPAVFRDSDRFSWHDPHLWRSGDEA
PGLFFVDAERVPCRHDDVFFPPSASFRVGLGPGASPVVRVRSISALGRTFTRDEDLAVFLASR
AGRRLRFHGPALSVGPEDCADPSGCVCGNAEAQPWICAALLQP

198/330

FIGURE 199

ATCGCATCAATTGGGAGTACCATCTTCCTCATGGGACCAGTGAAACAGCTGAAGCGAATGTT
TGAGCCTACTCGTTTGATTGCAACTATCATGGTGCTGTTGTGTTTTGCACTTACCCTGTGTT
CTGCCTTTTGGTGGCATAACAAGGGACTTGCACTTATCTTCTGCATTTTGCACTCTTTGGCA
TTGACGTGGTACAGCCTTTCCTTCATACCATTGCAAGGGATGCTGTGAAGAAGTGTTTTGC
CGTGTGTCTTGCATAATTTCATGGCCAGTTTTATGAAGCTTTGGAAGGCACTATGGACAGAAG
CTGGTGGACAGTTTTGTAACCTATCTTCGAAACCTCTGTCTTACAGACATGTGCCTTTTATCT
TGCAGCAATGTGTTGCTTGTGATTCGAACATTTGAGGGTTACTTTTGAAGCAACAATACAT
TCTCGAACCTGAATGTCAGTAGCACAGGATGAGAAGTGGGTCTGTATCTTGTGGAGTGGAA
TCTTCCTCATGTACCTGTTTCCTCTCTGGATGTTGTCCCACTGAATTCCCATGAATACAAAC
CTATTCAGCAACAGCAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAA

199/330

200/330

FIGURE 200

MGPVKQLKRMFEPTRLIATIMVLLCFALTLCSAFWWHNKGLALIFCILQSLALTWYSLSFIP
FARDAVKKCFVCLA

200/330

FIGURE 201

TTGAGCGCAGGTGAGCTCCTGCGCGTTCCGGGGGCGTTCCCTCCAGTCACCCTCCCGCCGTTACCCGCGGCGCGC
CCGAGGGAGTCTCCTCCAGACCCTCCCTCCCGTTGCTCCAACTAATACGGACTGAACGGATCGCTGCGAGGGT
GGGAGAGAAAATTAGGGGGAGAAAGGACAGAGAGCAACTACCATCCATAGCCAGATAGATTATCTTACACTG
AACTGATCAAGTACTTTGAAAATGACTTCGAAATTTATCTTGGTGTCTTCATACTTGCTGCACTGAGTCTTTC
AACCACCTTTTCTCTCCAAGTAGACCAGCAAAAGGTTCTACTAGTTTCTTTTGATGGATTCCGTTGGGATTACT
TATATAAAGTTCCAACGCCCCATTTTTCATTATATTATGAAATATGGTGTTCACGTGAAGCAAGTTACTAATGTT
TTTATTACAAAAACCTACCCTAACCATTATACTTTGGTAAGTGGCCTCTTTCAGAGAATCATGGGATTGTTGC
AAATGATATGTTTGATCCTATTCGGAACAAATCTTCTCCTTGGATCACATGAATATTTATGATTCCAAGTTTT
GGGAAGAAGCGACACCAATATGGATCACAACCAGAGGGCAGGACATACTAGTGGTGCAGCCATGTGGCCCGGA
ACAGATGTAAAAATACATAAGCGCTTTCTACTCATTACATGCCTTACAATGAGTCAGTTTCATTTGAAGATAG
AGTTGCCAAAAATTGTTGAATGGTTTACGTCAAAAGAGCCCATAAATCTTGGTCTTCTCTATTGGGAAGACCCTG
ATGACATGGGCCACCATTGGGACCTGACAGTCCGCTCATGGGCGCTGTCATTTTCAGATATTGACAAGAAGTTA
GGATATCTCATACAAATGCTGAAAAAGGCAAAGTTGTGGAACACTCTGAACCTAATCATCACAAGTGATCATGG
AATGACGCGAGTGCTCTGAGGAAAGGTTAATAGAAGTTGACCAGTACCTGGATAAAGACCATAACCTGATTG
ATCAATCTCCAGTAGCAGCCATCTTGCCAAAAGAAGGTAAATTTGATGAAGTCTATGAAGCACTAACTCACGCT
CATCCTAATCTTACTGTTTACAAAAAGAAGACGTTCCAGAAAGGTGGCATTACAAATACAACAGTCGAATTCA
ACCAATCATAGCAGTGGCTGATGAAGGGTGGCACATTTTACAGAATAAGTCAGATGACTTTCTGTTAGGCAACC
ACGGTTACGATAATGCGTTAGCAGATATGCATCCAATATTTTGTAGCCCATGGTCCTGCCTTCAGAAAGAATTTTC
TCAAAAGAAGCCATGAACTCCACAGATTTGTACCCACTACTATGCCACCTCCTCAATATCACTGCCATGCCACA
CAATGGATCATTCTGGAATGTCCAGGATCTGCTCAATTGAGCAATGCCAAGGGTGGTCCCTTATACACAGAGTA
CTATACTCCTCCCTGGTAGTGTTAAACCAGCAGAATATGACCAAGAGGGGTCTATACCCTTATTTTCATAGGGGTC
TCTCTTGGCAGCATTATAGTGATTGTATTTTGTAAATTTTCATTAAGCATTTAATTCACAGTCAAATACCTGC
CTTACAAGATATGCATGCTGAAATAGCTCAACCATTATTACAAGCCTAATGTTACTTTGAAGTGGATTTCGATA
TTGAAGTGGAGATTCCATAATTATGTCAGTGTTTAAAGGTTTCAAATTCCTGGGAAACAGTTCCAAACATCTGC
AGAAACCATTAAAGCAGTTACATATTTAGGTATACACACACACACACACACATACACACACACGACCAAAA
ATACTTACACCTGCAAAGGAATAAAGATGTGAGAGTATGTCTCCATTGTTCACTGTAGCATAGGGATAGATAAG
ATCCTGCTTTATTTGGACTTGGCGCAGATAATGTATATATTTAGCAACTTTGCACTATGTAAAGTACCTTATAT
ATTGCACTTTTAAATTTCTCTCCTGATGGGTACTTTAATTTGAAATGCACTTTATGGACAGTTATGTCTTATAAC
TTGATTGAAAATGACAACCTTTTGCACCCATGTCACAGAATACTTGTACGCATTGTTCAAACCTGAAGGAAATT
TCTAATAATCCCGAATAATGAACATAGAAATCTATCTCCATAAATTGAGAGAAGAAGAAGGTGATAAGTGTGTA
AAATTAAATGTGATAACCTTTGAACCTTGAATTTTGAGATGTATTCCCAACAGCAGAATGCAACTGTGGGCAT
TTCTTGTCTTATTTCTTTCCAGAGAACGTGGTTTTCTTTATTTTCCCTCAAAGAGAGTCAAATACTGACAG
ATTGCTTCTAAATATATTGTTTCTGTCAAAAATTATTGTGATTTCCCTGATGAGTCATATTACTGTGATTTTCA
TAATAATGAAGACACCATGAATATACTTTTCTTCTATATAGTTCAGCAATGGCCTGAATAGAAGCAACCAGGCA
CCATCTCAGCAATGTTTTCTCTTGTGTTGTAATTTATTGCTCCTTTGAAAATTAAATCACTATTAATTACATTAA
AAATCAAATTGGATAAAAAAAAAAAAAAAAAAAAA

FIGURE 202

MTSKFILVSFILAALSLSTTFSLQLDQQKVLVSFDGFRWDYLYKVPTPHFHYIMKYGVHVK
QVTNVFITKTYPNHYTLVTGLFAENHGIVANDMFDPIRNKSFSLDHMNIYDSKFWEEATPIW
ITNQRAGHTSGAAMWPGTDVKIHKRFPHTHYMPYNESVSFEDRVAKIVEWFTSKEPINLGLLY
WEDPDDMGHHLGPDSPLMGPVISDIDKKLGYLIQMLKKAKLWNTLNLIITSDHGMTQCSEER
LIELDQYLDKDHYTELIDQSPVAAILPKEGKFDEVYEALTHAHPNLTVYKKEDVPERWHYKYN
SRIQPIIAVADEGWHLQNKSDDFLLGNHGYDNALADMHPIFLAHGPAFRKNFSKEAMNSTD
LYPLLCHLLNITAMPHNGSFWNVQDLLNSAMPRVVPYTQSTILLPGSVKPAEYDQEGSYPYF
IGVSLGSIIVIVFFVIFIKHLIHSQIPALQDMHAEIAQPLLQA

Signal Peptide:

amino acids 1-22

Transmembrane Domain:

amino acids 429-452

N-glycosylation sites:

amino acids 101-104, 158-161, 292-295, 329-332, 362-365, 369-372, 382-385, 389-392

Somatomedin B Domain:

amino acids 69-85

Sulfatase protein Region:

amino acids 212-241

FIGURE 203

GGATTTTTGTGATCCGCGATTTCGCTCCACGGGCGGGACCTTTGTAAGTGCAGGGAGGCCAG
GACAGGGCCACCCTGCGGGGCGGGAGGCAGCCGGGGTGAGGGAGGTGAAGAAACCAAGACGC
AGAGAGGCCAAGCCCCCTTGCCTTGGGTACACAGCCAAAGGAGGCAGAGCCAGAACTCACAA
CCAGATCCAGAGGCAACAGGGAC**ATG**GCCACCTGGGACGAAAAGGCAGTCACCCGCAGGGCC
AAGGTGGCTCCCGCTGAGAGGATGAGCAAGTTCTTAAGGCACTTCACGGTCGTGGGAGACGA
CTACCATGCCTGGAACATCAACTACAAGAAATGGGAGAATGAAGAGGAGGAGGAGGAGGAGG
AGCAGCCACCACCCACACCAGTCTCAGGCGAGGAAGGCAGAGCTGCAGCCCCCTGACGTTGCC
CCTGCCCCCTGGCCCCGCACCCAGGGCCCCCCTTGACTTCAGGGGCATGTTGAGGAACTGTT
CAGCTCCCACAGGTTTCAGGTCATCATCATCTGCTTGGTGGTTCTGGATGCCCTCCTGGTGC
TTGCTGAGCTCATCCTGGACCTGAAGATCATCCAGCCCGACAAGAATAACTATGCTGCCATG
GTATTCCACTACATGAGCATCACCATCTTGGTCTTTTTTATGATGGAGATCATCTTTAAATT
ATTTGTCTTCCGCCTGAGTTCTTTCACCACAAGTTTGAGATCCTGGATGCCCGTCGTGGTGG
TGGTCTCATTCATCCTGGACATTGTCTCCTGTTCCAGGAGCACCAGTTTGAGGCTCTGGGC
CTGCTGATTCTGCTCCGGCTGTGGCGGGTGGCCCGGATCATCAATGGGATTATCATCTCAGT
TAAGACACGTTTCAGAACGGCAACTCTTAAGGTTAAAACAGATGAATGTACAATTGGCCGCCA
AGATTCAACACCTTGAGTTCAGCTGCTCTGAGAAGCCCCCTGGAC**TGA**TGAGTTTGCTGTATC
AACCTGTAAGGAGAAGCTCTCTCCGGATGGCTATGGGAATGAAAGAATCCGACTTCTACTCT
CACACAGCCACCGTGAAAGTCCTGGAGTAAATGTGCTGTGTACAGAAGAGAGAGAAGGAAG
CAGGCTGGCATGTTCACTGGGCTGGTGTACGACAGAGAACCTGACAGTCACTGGCCAGTTA
TCACTTCAGATTACAAATCACACAGAGCATCTGCCTGTTTTCAATCACAAGAGAACAAAACC
AAAATCTATAAAGATATTCTGAAAATATGACAGAATTTGACAAATAAAAGCATAAACGTGTA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 204

MATWDEKAVTRRAKVAPAERMSKFLRHFTVVGDDYHAWNINYKKWENEEEEEEEEQPPPTPV
SGEEGRAAAPDVAPAPGPAPRAPLDFRGMLRKLFSSHRFQVIIICLVVLDALLVLAELIDL
KIIQPDKNYAAMVFHYMSITILVFFMMEIIFKLFVFERLSSFTTSLRSWMPVVVVVSFILDI
VLLFQEHQFEALGLLILLRLWRVARIINGIIISVKTRSERQLRLKQMNVQLAAKIQHLEFS
CSEKPLD

204/330

CGGCTCGAGCTCGAGCCGAATCGGCTCGAGGGGAGCTGGAGCACCCAGCAGGCCGCCAACAT
GCTCTGTCTGTGCCTGTACGTGCCGGTCATCGGGGAAGCCCAGACCGAGTTCAGTACTTTG
AGTCGAAGGGGCTCCCTGCCGAGCTGAAGTCCATTTTCAAGCTCAGTGTCTTCATCCCCTCC
CAGGAATTCTCCACCTACCGCCAGTGGAAAGCAGAAAATTGTACAAGCTGGAGATAAAGGACCT
TGATGGGCAGCTAGACTTTGAAGAATTTGTCCATTATCTCCAAGATCATGAGAAGAGCTGA
GGCTGGTGTTTAAGATTTTGGACAAAAGAATGATGGACGCATTGACGCGCAGGAGATCATG
CAGTCCCTGCGGGACTTGGGAGTCAAGATATCTGAACAGCAGGCAGGAGAAAATTCTCAAGAG
CATGGATAAAAAACGGCACGATGACCATCGACTGGAACGAGTGGAGAGACTACCACCTCCTCC
ACCCCGTGGAAAACATCCCCGAGATCATCCTCTACTGGAAGCATTCCACGATCTTTGATGTG
GGTGAGAATCTAACGGTCCCGGATGAGTTCACAGTGGAGGAGAGGCAGACGGGGATGTGGTG
GAGACACCTGGTGGCAGGAGGTGGGGCAGGGCCGTATCCAGAACCTGCACGGCCCCCTGG
ACAGGCTCAAGTGCTCATGCGAGTCCATGCCTCCCGCAGCAACAACATGGGCATCGTTGGT
GGCTTCACTCAGATGATTCGAGAAGGAGGGGGCCAGGTCACCTCTGGCGGGGCAATGGCATCAA
CGTCTCTAAAATTGCCCCGAATCAGCCATCAAATTCATGGCCTATGAGCAGATCAAGCGCC
TTGTTGGTAGTGACCAGGAGACTCTGAGGATTACAGGAGGGCTTGTGGCAGGGTCTTGGCA
GGGGCCATCGCCAGAGCAGCATCTACCAATGGAGGTCTGAAGACCCGGATGGCGCTGCG
GAAGACAGGCCAGTACTCAGGAATGCTGACTGCGCCAGGAGGATCTGGCCAGAGAGGGGG
TGGCGCCTTCTACAAAGGCTATGTCCCCAACATGCTGGGCATCATCCCCCTATGCCGGCATC
GACCTTGCAGTCTACGAGACGCTCAAGAATGCCTGGCTGCAGCACTATGCAGTGAACAGCGC
GGACCCCGGCGTGTGTTGTCTGCTGCTGGCCTGTGGCACCATGTCCAGTACCTGTGGCCAGCTGG
CCAGCTACCCCCCTGGCCCTAGTCAAGACCCGGATGCAGGCGCAAGCCTCTATTGAGGCGCT
CCGGAGGTGACCATGAGCAGCCTCTTCAAACATATCTGCGGACCCAGGGGGCCCTCGGGCT
GTACAGGGGGCTGGCCCCCAACTTCATGAAGGTCATCCCAGCTGTGAGCATCAGCTACGTGG
TCTACGAGAACCTGAAGATCACCTTGGGCGTGCAGTTCGCGGT**TG**ACGGGGGGAGGGCCGCCCG
GCAGTGGACTCGCTGATCCTGGGCCGAGCCTGGGGTGTGCAGCCATCTCATTCTGTGAATG
TGCCAACACTAAGCTGTCTCGAGCCAAGCTGTGAAAACCTTAGACGCACCCCGAGGGAGGGT
GGGGAGAGCTGGCAGGCCAGGGCTTGTCTGCTGACCCAGCAGACCTCCTGTTGGTTCC
AGCGAAGACCACAGGCATTCCTTAGGGTCCAGGTCAGCAGGCTCCGGGCTCACATGTGTAA
GGACAGGACATTTTCTGCAGTGCCTGCCAATAGTGAGCTTGGAGCCTGGAGGCCGGCTTAGT
TCTTCCATTTTACCCTTGCAGCCAGCTGTTGGCCACGGCCCTGCCCTCTGGTCTGCCGTGC
ATCTCCCTGTGCCCTCTTGCTGCCTGCCTGTCTGCTGAGGTAAGGTGGGAGGAGGCTACAG
CCCACATCCCACCCCCCTCGTCCAATCCCATAATCCATGATGAAAGGTGAGGTACAGTGGCCT
CCCAGGCCTGACTTCCCAACCTACAGCATTGACGCCAATCTGGCTGTGAAGGAAGAGGAAAG
GATCTGGCCTTGTGGTCACTGGCATCTGAGCCCTGCTGATGGCTGGGGCTCTCGGGCATGCT
TGGGAGTGCAGGGGGCTCGGGCTGCCTGGCCTGGCTGCACAGAAGGCAAGTGTGGGGCTCA
TGGTGCTCTGAGCTGGCCTGGACCCCTGTGAGGATGGGCCCCACCTCAGAACCACACTGCTG
TCCCACTGTGGCATGAGGGCAGTGGAGCAGCATGTTTGAGGGCGAAGGCGAGAGCGTGTGT
GTGTTCTGGGGAGGGAAGGAAAAGGTGTTGGAGGCCTTAATTATGGACTGTTGGGAAAAGGG
TTTTGTCCAGAAGGACAAGCCGGACAAATGAGCGACTTCTGTGCTTCCAGAGGAAGACGAGG
GAGCAGGAGCTTGGCTGACTGCTCAGAGTCTGTTCTGACGCCCTGGGGGTTCTGTCCAACC
CCAGCAGGGGGCGCAGCGGGACAGCCCCACATTCACACTTGTGTCACTGCTTGAACCTATT
ATTTTGTATTTATTTGAACAGAGTTATGTCCTAATATTTTATAGATTTGTTTAAATTAATA
GCTTGTCAATTTCAAGTTCATTTTTTATTATCATATTTATGTTTATGGTTGATTGTACCTTCCC
AAGCCCGCCAGTGGGATGGGAGGAGGAGGAGAAGGGGGGCCCTTGGGCCGCTGCAGTCACAT
CTGTCCAGAGAAATTCCTTTTGGGACTGGAGGCAGAAAAGCGGCCAGAAGGCAGCAGCCCTG
GCTCCTTTTCTTTTGGCAGGTTGGGGAAGGGCTTGCCCCCAGCCTTAGGATTTAGGGTTTGA
CTGGGGGCGCTGGAGAGAGAGGAGGGAACCTCAATAACCTGAAGGTGGAATCCAGTTATTTG
CTGCGCTGCGAGGGGTTCTTTTACTCTTTTCTGAATGTCAAGGCAGTGAGGTGCCTCT
CACTGTGAATTTGTGGTGGGCGGGGGCTGGAGGAGAGGGTGGGGGGCTGGCTCCGTCCCTCC
CAGCCTTCTGCTGCCCTTGCTTAACAATGCCGGCCAACCTGGCGACCTCACGGTTGCACCTCC
ATTCCACCAGAATGACCTGATGAGGAAATCTTCAATAGGATGCAAAGATCAATGCAAAAATT
GTTATATATGAACATAATACTGGAGTCGTCAAAAAGCAAATTAAGAAAGAAATTGACGTTAG
AAGTTGTCAATTTAAAGCAGCCTTCTAATAAAGTTGTTTCAAAGCTGAAAAAATAAAAAA
AAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA

FIGURE 206

MLCLCLYVPVIGEAQTEFQYFESKGLPAELKSIFKLSVFIPSQEFSTYRQWKQKIVQAGDKD
LDGQLDFEEFVHYLQDHEKKLRLVFKILDKKNDGRIDAQEIMQSLRDLGVKISEQQAEEKILK
SMDKNGTMTIDWNEWRDYHLLHPVENIPEIILYWKHSTIFDVGENLTVPDEFTVEERQTGMW
WRHLVAGGGAGAVSRTCTAPLDRLKVLMOVHASRSNNMGIVGGFTQMIREGGARSLWRGNGI
NVLKIAPESAIFMAYEQIKRLVGSDQETLRIHERLVAGSLAGAIQSSIYPMEVLKTRMAL
RKTGQYSGMLDCARRILAREGVAAFYKGYVPNMLGIIPYAGIDLAVYETLKNWLQHYAVNS
ADPGVFVLLACGTMSSTCGQLASYPLALVRTRMQAASIEGAPEVTMSSLFKHILRTEGAFG
LYRGLAPNFMKVI PAVSISYVYENLKITLGVQSR

Important features:**Signal peptide:**

amino acids 1-16

Transmembrane domain:

amino acids 284-304, 339-360, 376-394

Mitochondrial energy transfer proteins signature.

amino acids 206-215, 300-309

N-glycosylation site.

amino acids 129-133, 169-173

Elongation Factor-hand calcium-binding protein.

amino acids 54-73, 85-104, 121-140

FIGURE 207

GGAAGGCAGCGGCAGCTCCACTCAGCCAGTACCCAGATACGCTGGGAACCTTCCCCAGCC**AT**
GGCTTCCCTGGGGCAGATCCTCTTCTGGAGCATAATTAGCATCATCATTATTCTGGCTGGAG
 CAATTGCACTCATCATTGGCTTTGGTATTTTCAGGGAGACACTCCATCACAGTCACTACTGTC
 GCCTCAGCTGGGAACATTGGGGAGGATGGAATCCTGAGCTGCACCTTTGAACCTGACATCAA
 ACTTTCTGATATCGTGATACAATGGCTGAAGGAAGGTGTTTTAGGCTTGGTCCATGAGTTCA
 AAGAAGGCAAAGATGAGCTGTCGGAGCAGGATGAAATGTTTCAGAGGCCGGACAGCAGTGTTT
 GCTGATCAAGTGATAGTTGGCAATGCCTCTTTGCGGCTGAAAAACGTGCAACTCACAGATGC
 TGGCACCTACAAATGTTATATCATCACTTCTAAAGGCAAGGGGAATGCTAACCTTGAGTATA
 AAAGTGGAGCCTTCAGCATGCCGGAAGTGAATGTGGACTATAATGCCAGCTCAGAGACCTTG
 CGGTGTGAGGCTCCCCGATGGTTCCCCCAGCCCACAGTGGTCTGGGCATCCCAAGTTGACCA
 GGGAGCCAACTTCTCGGAAGTCTCCAATACCAGCTTTGAGCTGAACTCTGAGAATGTGACCA
 TGAAGGTTGTGTCTGTGCTCTACAATGTTACGATCAACAACACATACTCCTGTATGATTGAA
 AATGACATTGCCAAAGCAACAGGGGATATCAAAGTGACAGAATCGGAGATCAAAAGGCGGAG
 TCACCTACAGCTGCTAAACTCAAAGGCTTCTCTGTGTGTCTCTTCTTTCTTTGCCATCAGCT
 GGGCACTTCTGCCTCTCAGCCCTTACCTGATGCTAAAA**TAA**TGTGCCTTGGCCACAAAAAAG
 CATGCAAAGTCATTGTTACAACAGGGATCTACAGAACTATTTTACCACCAGATATGACCTAG
 TTTTATATTTCTGGGAGGAAATGAATTCATATCTAGAAGTCTGGAGTGAGCAAACAAGAGCA
 AGAAACAAAAAGAAGCCAAAAGCAGAAGGCTCCAATATGAACAAGATAAATCTATCTTCAA
 GACATATTAGAAGTTGGGAAAATAATTCATGTGAACTAGACAAGTGTGTTAAGAGTGATAAG
 TAAAATGCACGTGGAGACAAGTGCATCCCCAGATCTCAGGGACCTCCCCCTGCCTGTCACCT
 GGGGAGTGAGAGGACAGGATAGTGCATGTTCTTTGTCTCTGAATTTTTAGTTATATGTGCTG
 TAATGTTGCTCTGAGGAAGCCCCTGGAAAGTCTATCCCAACATATCCACATCTTATATTCCA
 CAAATTAAGCTGTAGTATGTACCCTAAGACGCTGCTAATTGACTGCCACTTCGCAACTCAGG
 GGCGGCTGCATTTTAGTAATGGGTCAAATGATTCACCTTTTATGATGCTTCCAAAGGTGCCT
 TGGCTTCTCTTCCCAACTGACAAATGCCAAAGTTGAGAAAAATGATCATAATTTTAGCATAA
 ACAGAGCAGTCGGGGACACCGATTTTATAAATAAACTGAGCACCTTCTTTTAAACAAAAAA
 AA

FIGURE 208

MASLGQILFWSIISI IIIILAGAIALIIGFGISGRHSITVTTVASAGNIGEDGILSCTFEPDI
KLSDIVIQWLKEGVLGLVHEFKEGKDELSEQDEMFRGRTAVFADQVIVGNASLRLKNVQLTD
AGTYKCYIIITSKGKGANLEYKTGAFSMPEVNVVDYNASSETLRCEAPRWFPQPTVVWASQVD
QGANFSEVSNTSFELNSENVTMKVVSVLNVNTINNTYSCMIENDIAKATGDIKVTESEIKRR
SHLQLLNSKASLCVSSFFAISWALLPLSPYLMLK

208

[illegible]

210/330

FIGURE 210

MAASLGQVLALVLVAALWGGTQPLLKRASAGLQRVHEPTWAQQLQEMKTLFLNTEYLMPFL
LNQCGSLLYYLTLASTDLTLAVPICNSLAIIFTLIVGKALGEDIGGKRKLDYCECGTQLCGS
RHTCVSSFPEPISPEWVRTRPFPILPFPLQLFCFLVAIRVPFPWTVWRKTEAGVWD

210/330

FIGURE 211

CTTCTGTAGGACAGTCACCAGGCCAGATCCAGAAGCCTCTCTAGGCTCCAGCTTTCTCTGTG
GAAGATGACAGCAATTATAGCAGGACCCTGCCAGGCTGTGCGAAAAGATTCCGCAATAAACT
TTGCCAGTGGGAAGTACCTAGTGAAACGGCCTAAGATGCCACTTCTTCTCATGTCCCAGGCT
TGAGGCCCTGTGGTCCCCATCCTTGGGAGAAGTCAGCTCCAGCACCATGAAGGGCATCCTCG
TTGCTGGTATCACTGCAGTGCTTGTGTCAGCTGTAGAATCTCTGAGCTGCGTGCAGTGTAAT
TCATGGGAAAAATCCTGTGTCAACAGCATTGCCTCTGAATGTCCCTCACATGCCAACACCAG
CTGTATCAGCTCCTCAGCCAGCTCCTCTCTAGAGACACCAGTCAGATTATAACCAGAATATGT
TCTGCTCAGCGGAGAACTGCAGTGAGGAGACACACATTACAGCCTTCACTGTCCACGTGTCT
GCTGAAGAACACTTTTCATTTTGTAAGCCAGTGCTGCCAAGGAAAGGAATGCAGCAACACCAG
CGATGCCCTGGACCCCTCCCCTGAAGAACGTGTCCAGCAACGCAGAGTGCCCTGCTTGTTATG
AATCTAATGGAACCTCCTGTGTCGTGGGAAGCCCTGGAAATGCTATGAAGAAGAACAGTGTC
TTTCTAGTTGCAGAACTTAAGAATGACATTGAGTCTAAGAGTCTCGTGCTGAAAGGCTGTTC
CAACGTCAGTAACGCCACCTGTCAGTTCCTGTCTGGTGAAAACAAGACTCTTGGAGGAGTCA
TCTTTCGAAAGTTTGAGTGTCAAATGTAAACAGCTTAACCCCCACGTCTGCACCAACCACT
TCCCACAACGTGGGCTCCAAAGCTTCCCTCTACCTCTTGGCCCTTGCCAGCCTCCTTCTTCG
GGGACTGCTGCCCCTGAGGTCCTGGGGCTGCACTTTGCCCAGCACCCCATTTCTGCTTCTCTG
AGGTCCAGAGCACCCCTGCGGTGCTGACACCCTCTTTCCTTGCTCTGCCCCGTTTAACTGC
CCAGTAAGTGGGAGTCACAGGTCTCCAGGCAATGCCGACAGCTGCCTTGTTCTTCATTATTA
AAGCACTGGTTCATTCAC TGCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 212

MKGILVAGITAVLVAAVESLSCVQCNSWEKSCVNSIASECPSHANTSCISSSASSSLETPVR
LYQNMFCSAENCSEETHITAFTVHVSAEEHFHFVSQCCQGKECSNTSDALDPPLKNVSSNAE
CPACYESNGTSCRGKPWKCYEEEQCVFLVAELKNDIESKSLVLKGCSNVSNATCQFLSGENK
TLGGVIFRKFEKANVNSLTPTSAPTTSNVGSKASLYLLALASLLLRGLLP

212/330

FIGURE 213

GGCCTCGGTTCAAACGACCCGGTGGGTCTACAGCGGAAGGGAGGGAGCGAAGGTAGGAGGCA
GGGCTTGCCTCACTGGCCACCTCCCAACCCCAAGAGCCCAGCCCC**ATG**GTCCCCGCCGCCG
GCGCGCTGCTGTGGGTCTGCTGCTGAATCTGGGTCCCCGGGCGGCGGGGGCCCAAGGCCTG
ACCCAGACTCCGACCGAAATGCAGCGGGTCAGTTTACGCTTTGGGGGGCCCCATGACCCGCAG
CTACCGGAGCACCGCCCGGACTGGTCTTCCCCGGAAGACAAGGATAATCCTAGAGGACGAGA
ATGATGCCATGGCCGACGCCGACCGCCTGGCTGGACCAGCGGCTGCCGAGCTCTTGGCCGCC
ACGGTGTCCACCGGCTTTAGCCGGTCGTCCGCCATTAACGAGGAGGATGGGTCTTCAGAAGA
GGGGGTGTGATTAATGCCGGAAAGGATAGCACCAGCAGAGAGCTTCCAGTGCAGACTCCCA
ATACAGCGGGGAGTTCCAGCACGAGGTTTATAGCCAATAGTCAGGAGCCTGAAATCAGGCTG
ACTTCAAGCCTGCCGCGCTCCCCGGGAGGTCTACTGAGGACCTGCCAGGCTCGCAGGCCAC
CCTGAGCCAGTGGTCCACACCTGGGTCTACCCCGAGCCGGTGGCCGTCACCTCACCCACAG
CCATGCCATCTCCTGAGGATCTGCGGCTGGTGCTGATGCCCTGGGGCCCGTGGCACTGCCAC
TGCAAGTCGGGCACCATGAGCCGGAGCCGGTCTGGGAAGCTGCACGGCCTTTCCGGGCGCCT
TCGAGTTGGGGCGCTGAGCCAGCTCCGCACGGAGCACAAGCCTTGACCTATCAACAATGTC
CCTGCAACCGACTTCGGGAAGAGTGCCCCCTGGACACAAGTCTCTGTACTGACACCAACTGT
GCCTCTCAGAGCACCACCAGTACCAGGACCACCACTACCCCCTTCCCCACCATCCACCTCAG
AAGCAGTCCCAGCCTGCCACCCGCCAGCCCCTGCCAGCCCTGGCTTTTTTGGAAACGGGTCA
GGATTGGCCTGGAGGATATTTGGAATAGCCTCTCTTCAGTGTTACAGAGATGCAACCAATA
GACAGAAACCAGAGG**TAA**TGGCCACTTCATCCACATGAGGAGATGTCAGTATCTCAACCTCT
CTTGCCCTTTCAATCCTAGCACCCACTAGATATTTTGTAGTACAGAAAAACAAAACCTGGAAAA
CACAA

FIGURE 214

MVPAAGALLWVLLLNLGPRAGAAGLTQTPTEMQRVSLRFGGPMTRSYRSTARTGLPRKTRI
ILEDENDAMADADRLAGPAAAELLAATVSTGFSRSSAINEEDGSSEEGVINAGKDSTSREL
PSATPNTAGSSSTRFIANSQEPEIRLTSSLPRSPGRSTEDLPGSQATLSQWSTPGSTPSRWP
SPSPTAMPSPEDLRLVLMWPWGPWHCHCKSGTMSRSRSGKLHGLSGRLRVGALSQLRTEHKPC
TYQQPCPNRLREECPLDTSICTDTNCASQSTTSTRTTTTPFPTIHLRSSPSLPPASPCPALA
FWKRVRIGLEDIWNSLSSVFTEMQPIDRNQR

214/330

FIGURE 216

MAGLAARLVLLAGAAALASGSQGDREPVYRDCVLQCEEQNCSGGALNHFRSRQPIYMSLAGW
TCRDDCKYECMWVTVGLYLQEGHKVPQFHGKWPFSSRFLFFQEPASAVASFLNGLASLVMLCR
YRTFVPASSPMYHTCVAFAWVSLNAFWSTVFHTRDSDLTEKMDYFCASTVILHSIYLCCVR
TVGLQHPAVVSAFRALLLLMLTVHVSYSLSLIRFDYGYNLVANVAIGLVNVVWWLAWCLWNQR
RLPHVRKCVVVVLLQLGLSLLELLDFPPLFWVLDAHAIWHISTIPVHVLFSSFLEDDSLYLL
KESEDKFKLD

Important features:**Signal peptide:**

amino acids 1-20

Transmembrane domains:

amino acids 105-123, 138-156, 169-185, 193-209, 221-240, 256-272

N-glycosylation site.

amino acids 40-44

N-myristoylation site.

amino acids 43-49

CUB domain proteins profile.

amino acids 285-302

Amiloride-sensitive sodium channels proteins.

amino acids 162-186

FIGURE 217

[illegible]

FIGURE 218

MAPQSLPSSRMAPLGMLLGLLMAACFTFCLSHQNLKEFALTNPEKSSTKETERKETKAEEL
DAEVLEVFPHEWQALQPGQAVPAGSHVRLNLQTGEREAKLQYEDKFRNNLKGKRLDINTN
TYTSQDLKSALAKFKEGAEMESSKEDKARQAEVKRLFRPIEELKKDFDELNVVIETDMQIMV
RLINKFNSSSSSLEEKIAALFDLEYVHQM DNAQDLLSF GGLQVVINGLNSTEPLVKEYAAF
VLGAAFSSNPVKVQVEAIEGGALQKLLVILATEQPLTAKKKVLFALCSLLRHFPYAQRQFLKL
GGLQVLR TLVQEKGTEVLAVRVV TLLYDLVTEKMFAEEEAELTQEMSPEKLQQYRQVHLLPG
LWEQGWCEITAHLLALPEHDAREKVLQTLGVLLTTCRDRYRQDPQLGRTLASLQAEYQVLAS
LELQDGEDEGYFQELLGSVNSLLKELR

Important features:**Signal peptide:**

amino acids 1-29

Hypothetical YJL126w/YLR351c/yhcX family protein.

amino acids 364-373

N-glycosylation site.

amino acids 193-197, 236-240

N-myristoylation site.

amino acids 15-21, 19-25, 234-240, 251-257, 402-408, 451-457

Homologous region SLS1 protein.

amino acids 68-340

FIGURE 219

TTCGGCTTCCGTAGAGGAAGTGGCGCGGACCTTCATTTGGGGTTTCGGTTCCCCCCTTCCC
CTTCCCCGGGGTCTGGGGGTGACATTGCACCGCGCCCCCTCGTGGGGTCGCGTTGCCACCCCA
CGCGGACTCCCCAGCTGGCGCGCCCCCTCCCATTTGCCTGTCTGGTCAGGCCCCCACCCTT
TTCCACCTGACCAGCC**ATG**GGGGCTGCGGTGTTTTTCGGCTGCACTTTCGTCGCGTTTCGGC
CCGGCCTTCGCGCTTTTTCTTGATCACTGTGGCTGGGGACCCGCTTCGCGTTATCATCCTGGT
CGCAGGGGCATTTTTCTGGCTGGTCTCCCTGCTCCTGGCCTCTGTGGTCTGGTTCATCTTGG
TCCATGTGACCGACCGGTAGATGCCCGGCTCCAGTACGGCCTCCTGATTTTTGGTGCTGCT
GTCTCTGTCTTCTACAGGAGGTGTTCCGCTTTGCCTACTACAAGCTGCTTAAGAAGGCAGA
TGAAGGGTTAGCATCGCTGAGTGAGGACGGAAGATCACCCATCTCCATCCGCCAGATGGCCT
ATGTTTCTGGTCTCTCCTTCGGTATCATCAGTGGTGTCTTCTCTGTTATCAATATTTGGCT
GATGCACTTGGGCCAGGTGTGGTTGGGATCCATGGAGACTCACCTATTACTTCCTGACTTC
AGCCTTTCTGACAGCAGCCATTATCCTGCTCCATACCTTTTGGGGAGTTGTGTTCTTTGATG
CCTGTGAGAGGAGACGGTACTGGGCTTTGGGCCTGGTGGTTGGGAGTCACCTACTGACATCG
GGACTGACATTCTGAACCCCTGGTATGAGGCCAGCCTGCTGCCCATCTATGCAGTCACTGT
TTCCATGGGGCTCTGGGCCTTCATCACAGCTGGAGGGTCCCTCCGAAGTATTCAGCGCAGCC
TCTTGTGTAAGGACT**TGA**CTACCTGGACTGATCGCCTGACAGATCCCACCTGCCTGTCCACTG
CCCATGACTGAGCCCAGCCCCAGCCGGGTCCATTGCCACATTCTCTGTCTCCTTCTCGTC
GGTCTACCCCACTACCTCCAGGGTTTTGCTTTGTCTTTTGTGACCGTTAGTCTCTAAGCTT
TACCAGGAGCAGCCTGGGTTTACGCCAGTCAGTGACTGGTGGGTTTTGAATCTGCACTTATCCC
CACCACCTGGGGACCCCTTGTGTGTCCAGGACTCCCCCTGTGTGCTGCTCTGCTCTCAC
CCTGCCCAAGACTCACCTCCCTTCCCCTCTGCAGGCCGACGGCAGGAGGACAGTCGGGTGAT
GGTGTATTCTGCCCTGCGCATCCCACCCGAGGACTGAGGGAACCTAGGGGGGACCCCTGGGC
CTGGGGTGCCCTCCTGATGTCTCGCCCTGTATTTCTCCATCTCCAGTTCTGGACAGTGCAG
GTTGCCAAGAAAAGGGACCTAGTTTAGCCATTGCCCTGGAGATGAAATTAATGGAGGCTCAA
GGATAGATGAGCTCTGAGTTTCTCAGTACTCCCTCAAGACTGGACATCTTGGTCTTTTTCTC
AGGCCTGAGGGGGAACCATTTTTGGTGTGATAAATACCCTAAACTGCCTTTTTTTCTTTTTT
GAGGTGGGGGGAGGGAGGAGGTATATTGGAACCTCTTCTAACCTCCTTGGGCTATATTTCTC
TCCTCGAGTTGCTCCTCATGGCTGGGCTCATTTTCGGTCCCTTTCTCCTTGGTCCCAGACCTT
GGGGGAAAGGAAGGAAGTGCATGTTTGGGAACTGGCATTACTGGAACATAATGGTTTTAACCT
CCTTAACCACCAGCATCCCTCCTCTCCCCAAGGTGAAGTGGAGGGTGCTGTGGTGAGCTGGC
CACTCCAGAGCTGCAGTGCCACTGGAGGAGTCAGACTACCATGACATCGTAGGGAAGGAGGG
GAGATTTTTTTGTAGTTTTTAATTGGGGTGTGGGAGGGGCGGGGAGGTTTTCTATAAACTGT
ATCATTTTCTGCTGAGGGTGGAGTGTCCCATCCTTTTAATCAAGGTGATTGTGATTTTGACT
AATAAAAAAGAATTTGTAAAAA
AAA

220/330

FIGURE 220

MGAAVFFGCTFVAFGPAFALFLITVAGDPLRVIIILVAGAFFWLVSLLLASVVWFILVHVTDR
SDARLQYGLLIFGAAVSVLLQEVRFRFAYYKLLKKADEGLASLSEDGRSPISIRQMAYVSGLS
FGIISGVFSVINILADALGPGVVGIIHGDSPPYYFLTSAFLTAAIILLHTFWGVVFFDACERRR
YWALGLVVGSHLLTSGLTFLNPWYEASLLPIYAVTVSMGLWAFITAGGSLRSIQRSLLCKD

220/330

FIGURE 221

AAGCTGGTTTAAAGGAAGCAGAGGAGGGTTAGATTTCGTTGAGTGAGGACGGAAGATCAACCCA
TTTCCATTCCGCCAGATGGCCTATGTTTCTGGTCTCTCCCTTCGGNATCATCAGTGGTGTNT
TNTCTGTTATCAATATTTTGGCTGATGCANTTGGGCCAGGTGTGGTTGGGATCCATGGAGAC
TCACCCTATTANTTCCTGANTTCAGCCTTTNTGACAGCAGCCATTATCCTGCTC

Figure 221

FIGURE 222

GACCGACCGTTCAGATGCCCCGGTTCCAGTACGGCTTCCTGATTTTTGGTGCTGCTGTNTCTG
TCCTTCTACAGGAGGTGTTCCGCTTTGCCTANTACAAGCTGCTTAAGAAGGCAGATGAGGGG
TTAGCATNGCTGAGTGAGGACGGAAGATCACCCATTTCCATCCGCCAGATGGCCTATGTTTN
TGGTNTTTCCTTCGGTATCATCAGTGGTGTNTCTGTTATCAATATTTTGGNTGATGCAN
TTGGGCCAGGTGTGGTTGGGATCCATGGAGANTCACCTATTAATTCCTGAATTCAGCCTTT
NTGACAGCAGCCATTATCCTGNTCCATACCTTTTGGGGAGTTGTGTTTTTTGATGCCTGTGA
GAGGAG

222/330

FIGURE 223

NGTTGGAGAAAGTGGCGCGGACNTTCATTTGGGGTTTCGGTTTCCCCCTTTCCCTTTCCCCG
GGGTCTGGGGTGACATTGCACGGGCCCCCTCGTGGGGTCGCGTTGCCACCCACGCGGACTCC
CCAGNTGGNGCGCCCTTCCCATTTGCCTGTCCTGGTCAGGCCCCCACCCTTCCCACNTG
ACCAGCCATGGGGGTGCGGTGTTTTTCGGCTGCACTTTCGTCGCGTTCGGCCCGGCCTTCG
CGCTTTTCTTGATCACTGTGGCTGGGGACCCGCTTCGCGTTATCATCCTGGTCGCAGGGGCA
TTTTTCTGGCTGGTCTCCCTGCTCCTGGCCTCTGTGGTCTGGTTCATCTTGGTCCATGTGAC
CGACCGGTCAGATGCCCCGGCTCCAGTACGGCCTCCTGATTTTTGGTGCTGCTGTCTCTGTCC
TTCTACAGGAGGTGTTCCGCTTTGCCTACTACAAGCTGCTTAAGAAGGCAGATGAGGGGTTA
GCATCGCTGAGTGAGGACGGAAGATCACCCATCTCCATCCGCCAGATGGCCTATGTTTCTGG
TCTCTCCTTCGGTATCATCAGTGGTGTCTTCTCTGTTATCAATATTTGGCTGATGCACTTG
GGCCAGGTGTGGTTGGGATCCATGGAGACTCACCC

FIGURE 224

GTAAAAGAAAGTGGCCGGACCTTCATTGGGGTTTCGGTTCCCCCCTTCCCNTTCCCCGGGG
TCTGGGGGTGACATTGCACCGCGCCCNCTCGTGGGGTCGCGTTGCCACCCACGCGGACTCCC
CAGNTGGCGCGCCCTCCCATTTCCTGTCCTGGTCAGGCCCCACCCCCCTTCCCACCTGA
CCAGCCATGGGGGCTGCGGTGTTTTTCGGGCTGCACTTTCGTCGCGTTCGGGCCCGGCCTTC
GCGCTTTTCTTGATCACTGTGGCTGGGGACCCGCTTCGCGTTATCATCCTGGTCGCAGGGGC
ATTTTTCTGGCTGGTCTCCCTGCTCCTGGCCTCTGTGGTCTGGTTCATCTTGGTCCATGTGA
CCGACCGGTGAGATGCCCCGGCTCCAGTACGGCCTCCTGATTTTTGGTGCTGCTGTCTCTGTC
CTTCTACAGGAGGTGTTCCGCTTTGCCTACTACAAGCTGCTTAAGAAGGCAGATGAGGGGTT
AGCATCGCTGAGTGAGGACGGAAGATCACCCATCTCCATCCGCCAGATGGCCTATGTTTCTG
GTCTCTCCTTCGGTATCATCAGTGGTGTCTTCTCTGTTATCAATATTTTGGCTGATGCACTT
GGGCCAGGTGTGGTTGGGATCCATGGAGAC

FIGURE 225

CCCCCAGGGAGCAGTGGGTGGTTATAACTCAGGCCCGGTGCCCAGAGCCCAGGAGGAGGCAG
TGGCCAGGAAGGCACAGGCCTGAGAAGTCTGCGGCTGAGCTGGGAGCAAATCCCCACCCCC
TACCTGGGGGACAGGGCAAGTGAGACCTGGTGAGGGTGGCTCAGCAGGCAGGGAAGGAGAGG
TGTCTGTGCGTCCTGCACCCACATCTTTCTCTGTCCCCTCCTTGCCCTGTCTGGAGGCTGCT
AGACTCCTATCTTCTGAATTCTATAGTGCCTGGGTCTCAGCGCAGTGCCGATGGTGGCCCGT
CCTTGTGGTTCCTCTCTACCTGGGGAAATAAGGTGCAGCGGCC**ATGG**CTACAGCAAGACCCC
CCTGGATGTGGGTGCTCTGTGCTCTGATCACAGCCTTGCTTCTGGGGGTACAGAGCATGTT
CTCGCCAACAATGATGTTTTCTGTGACCACCCCTCTAACACCGTGCCCTCTGGGAGCAACCA
GGACCTGGGAGCTGGGGCCGGGGAAGACGCCCCGGTCCGATGACAGCAGCAGCCGCATCATCA
ATGGATCCGACTGCGATATGCACACCCAGCCGTGGCAGGCCGCGCTGTTGCTAAGGCCCAAC
CAGCTCTACTGCGGGGCGGTGTTGGTGCATCCACAGTGGCTGCTCACGGCCGCCCACTGCAG
GAAGAAAGTTTTAGAGTCCGTCTCGGCCACTACTCCCTGTCACCAGTTTATGAATCTGGGC
AGCAGATGTTCCAGGGGGTCAAATCCATCCCCACCCCTGGCTACTCCCACCCCTGGCCACTCT
AACGACCTCATGCTCATCAAACCTGAACAGAAGAATTCGTCCCCTAAAGATGTCAGACCCAT
CAACGTCTCCTCTCATTTGTCCCTCTGCTGGGACAAAGTGCTTGGTGTCTGGCTGGGGGACAA
CCAAGAGCCCCCAAGTGCACTTCCCTAAGGTCCTCCAGTGCTTGAATATCAGCGTGCTAAGT
CAGAAAAGGTGCGAGGATGCTTACCCGAGACAGATAGATGACACCATGTTCTGCGCCGGTGA
CAAAGCAGGTAGAGACTCCTGCCAGGGTGATTCTGGGGGGCCTGTGGTCTGCAATGGCTCCC
TGCAGGGACTCGTGTCCTGGGGAGATTACCCTTGTTGCCCGGCCCAACAGACCGGGTGTCTAC
ACGAACCTCTGCAAGTTCACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCC**TGAG**TGCAT
CCCAGGACTCAGCACACCGGCATCCCCACCTGCTGCAGGGACAGCCCTGACACTCCTTTCAG
ACCCTCATTCTTCCCAGAGATGTTGAGAATGTTTCATCTCTCCAGCCCCTGACCCCATGTCT
CCTGGACTCAGGGTCTGCTTCCCCCACATTGGGCTGACCGTGTCTCTCTAGTTGAACCCTGG
GAACAATTTCCAAAACCTGTCCAGGGCGGGGGTTGCGTCTCAATCTCCCTGGGGCACTTTCAT
CCTCAAGCTCAGGGCCCATCCCTTCTCTGCAGCTCTGACCCAAATTTAGTCCCAGAAATAAA
CTGAGAAGTGGAATAAAAAA

FIGURE 226

MATARPPMMWVLCALITALLGVTEHVLANNDVSCDHPSNTVPSGSNQDLGAGAGEDARSDD
SSSRIINGSDCDMHTQPWQAALLLRPNQLYCGAVLVHPQWLLTAAHCRKKVFRVRLGHYSLS
PVYESGQQMFQGVKSIPHPGYSHPGHSNDLMLIKLNRRIRPTKDVRPINVSSHCP SAGTKCL
VSGWGTTKSPQVHFPKVLQCLNISVLSQKRCEDAYPRQIDDTMFCAGDKAGR DSCQGD SGGP
VVCNGSLQGLVSWGDYPCARPNRPGVYTNLCKFTKWIQETIQANS

Downloaded from www.sciencedirect.com

FIGURE 227

ATGGTCAACGACCGGTGGAAGACCATGGGCGGCGCTGCCCAACTTGAGGACCGGGCCGCGGA
 CAAGCCGCAGCGGCCGAGCTGCGGGCTACGTGCTGTGCACCGTGCTGCTGGCCCTGGCTGTGC
 TGCTGGCTGTAGCTGTACCGGTGCCGTGCTCTTCTGAACCACGCCCACGCGCCGGGCACG
 GCGCCCCACCTGTCGTCAGCACTGGGGCTGCCAGCGCCAACAGCGCCCTGGTCACTGTGGA
 AAGGGCGGACAGCTCGCACCTCAGCATCCTCATTGACCCGCGCTGCCCGACCTCACCGACA
 GCTTCGCACGCCTGGAGAGCGCCCAGGCCTCGGTGCTGCAGGCGCTGACAGAGCACCAGGCC
 CAGCCACGGCTGGTGGGCGACCAGGAGCAGGAGCTGCTGGACACGCTGGCCGACCAGCTGCC
 CCGGCTGCTGGCCGAGCCTCAGAGCTGCAGACGGAGTGCATGGGGCTGCGGAAGGGGCATG
 GCACGCTGGGCCAGGGCCTCAGCGCCCTGCAGAGTGAGCAGGGCCGCCTCATCCAGCTTCTC
 TCTGAGAGCCAGGGCCACATGGCTCACCTGGTGAACCTCCGTACGCGACATCCTGGATGCCCT
 GCAGAGGGACCGGGGGCTGGGCGGGCCCCGCAACAAGGCCGACCTTCAGAGAGCGCCTGCC
 GGGGAACCCGGCCCCGGGGCTGTGCCACTGGCTCCCGCCCCGAGACTGTCTGGACGTCCTC
 CTAAGCGGACAGCAGGACGATGGCGTCTACTCTGTCTTTCCACCCACTACCCGGCCGGCTT
 CCAGGTGTACTGTGACATGCGCACGGACGGCGGGCGGTGGACGGTGTTCAGCGCCGGGAGG
 ACGGCTCCGTGAACCTTCTCCGGGGCTGGGACGCGTACCGAGACGGCTTTGGCAGGCTCACC
 GGGGAGCACTGGCTAGGGCTCAAGAGGATCCACGCCCTGACCACACAGGCTGCCTACGAGCT
 GCACGTGGACCTGGAGGACTTTGAGAATGGCAGGGCCTATGCCCGCTACGGGAGCTTCGGCG
 TGGGCTTGTCTCCGTGGACCCCTGAGGAAGACGGGTACCCGCTCACCGTGGCTGACTATTCC
 GGCACTGCAGGCGACTCCCTCCTGAAGCACAGCGGCATGAGGTTACCAACCAAGGACCGTGA
 CAGCGACCATTCAGAGAACAACCTGTGCCGCTTCTACCGCGGTGCCTGGTGGTACCGCAACT
 GCCACACGTCCAACCTCAATGGGCAGTACCTGCGCGGTGCGCACGCCTCCTATGCCGACGGC
 GTGGAGTGGTCTCCTGGACCGGTGGCAGTACTCACTCAAGTTCTCTGAGATGAAGATCCG
 GCCGGTCCGGGAGGACCGCT**AGACTGGTGACCTTGTCTTGGCCCTGCTGGTCCCTGT**CGC
 CCCATCCCCGACCCACCTCACTCTTTCTGAATGTTCTCCACCCACCTGTGCCTGGCGGAC
 CCACTCTCCAGTAGGGAGGGGCCGGGCCATCCCTGACACGAAGCTCCCTGGGCCGGTGAAGT
 CACACATCGCCTTCTCGCCGTCCCCACCCCTCCATTTGGCAGCTCACTGATCTCTTGCCCTC
 TGCTGATGGGGCTGGCAAACCTTGACGACCCCAACTCCTGCCTGCCCCACTGTGACTCCGG
 TGCTGTTTGGCGTCCCTGGCCAGGATGGTGGAGTCTGCCCCAGGCACCCCTCTGCCCTGCC
 GGCCAAATACCCGGCATTATGGGGACAGAGAGCAGGGGGCAGACAGCACCCTGGAGTCCTC
 CTAGCAGATCGTGGGGAATGTGAGGTCTCTCTGAGGTGAGGTCTGAGGCCAGTATCCTCCAG
 CCCTCCCAATGCCAACCCCAACCCCGTTTCCCTGGTGCCAGAGAACCCACCTCTCCCCAA
 GGGCCTCAGCCTGGCTGTGGGCTGGGTGGCCCCATCCTACCAGGCCCTGAGGTGAGGATGGG
 GAGCTGCTGCCTTTGGGGACCCACGCTCCAAGGCTGAGACCAGTTCCTGGAGGCCACCCAC
 CCTGTGCCCCGGCAGGCCTGGGGTCTGCAGTCTCTTACCTGCTGTGCCACCTGCTCTCTG
 TCTCAAATGAGGCCCCAACCCATCCCCACCCAGCTCCCGGCCGTCTCTACCTGGGGCAGC
 CGGGGCTGCCATCCCATTCTCTGCTCTGGAAGGTGGGTGGGGCCCTGCACCGTGGGGCT
 GGACTGCGCTAATGGGAAGCTCTTGGTTTTCTGGGCTGGGGCCTAGGCAGGGCTGGGATGAG
 GCTTGTAACCCCCACCAACCAATTTCCAGGGACTCCAGGGTCTGAGGCCTCCAGGAGG
 GCCTTGGGGGTGATGACCCCTTCCCTGAGGTGGCTGTCTCCATGAGGAGGCCAACCTTGCC
 ATTGACCGTGGCCACCTGGACCCAGGCCAGGCCCGGCCGCGAGTGGTCAAGGGACAGGGA
 CCACCTCACCGGGCAAATGGGGTGGGGGGGACTGGGGCACCAGACCAGGCACCACCTGGACA
 CTTTCTTGTGAATCCTCCCAACACCCAGCAGCTGTGCATCCCCACTCCTTGTGTGCACACA
 TGCAGAGGTGAGACCCGACGGCTCCCAGGACCAGCAGCCACAAGGGCAGGGCTGGAGCCGGG
 TCCTCAGCTGTCTGCTCAGCAGCCCTGGACCCGCGTGCCTTACGTGAGGCCAGATGCAGGG
 CGGCTTTTCCAAGGCCTCCTGATGGGGGCTCCGAAAGGGCTGGAGTCAGCCTTGGGGAGCT
 GCCTAGCAGCCTCTCCTCGGGCAGGAGGGGAGGTGGCTTCTTCCAAGGACACCCGATGGCA
 GGTGCCTAGGGGGTGTGGGGTTCCGTTCTCCCTCCCTCCCACTGAAGTTTGTGCTTAAAA
 AACATAAATTTGACTTGGCACCCTGGGGGTGGTGGGAGAGGCCGTGTGACCTGGCTCTC
 TGTCCAGTGCCACCAGGTATCCACATGCGCAG

FIGURE 228

MVNDRWKTMGGAAQLEDRPRDKPQRPSCGYVLCTVLLALAVLLAVAVTGAVLFLNHAHAPGT
APPPVVSTGAASANSALVTVERADSSHLSILIDPRCPDLTDSFARLESAQASVLQALTEHQA
QPRLVGDQEQELLDTLADQLPRLRARASELQTECMGLRKGHGTLGQGLSALQSEQGRLIQLL
SESQGHMAHLVNSVSDILDALQRDRGLGRPRNKADLQRAPARGTRPRGCATGSRPRDCLDVL
LSGQQDDGVYSVFPTHYPAGFQVYCDMRTDGGGWTVFQRRDGSVNFFRGWDAYRDGFGRLT
GEHWLGLKRIHALTTQAAYELHVDLEDFENGTAARYGSFGVGLFSVDPEEDGYPLTVADYS
GTAGDSLLKHSGMRFTTKDRSDHSENNCAAFYRGAWWYRNCHTSNLNGQYLRGAHASADG
VEWSSWTGWQYSLKFSEMKIRPVREDR

228/330

230/330

FIGURE 230

MQAKYSSTRDMLDDDGDTTMSLHSQASATTRHPEPRRTEHRAPSSTWRPVALTLLTLCVL
IGLAALGLLFFQYYQLSNTGQDTISQMEERLGNTSQELQSLQVQNIKLAGSLQHVAEKL
LYNKAGAHRCSPCTEQWKWHGDNCYQFYKDSKSWEDCKYFCLSENSTMLKINKQEDLE
QSYSEFFYSYWTGLLRPDSGKAWLWMDGTPFTSELFHIIIDVTSPRSRDCVAILNGMI
FSKDCKELKRCVCERRAGMVKPESLHVPPELGE

230/330

FIGURE 233

MRASLLLSVLRPAGPVAVGISLGFTLSLLSVTWVEEPCGPGPPQPGDSELP PRGNTNAARRP
NSVQPGAEREKPGAGEGAGENWEPRVLPYHPAQPGQAAKKAVRTRYISTELGIRQRLLVAVL
TSQTTLPTLGVAVNRTLGHRLERVVFLTGARGRRAPPGMAVVTLGEERPIGHLHLALRHLE
QHGDDEFDWFFLVPDTTYTEAHGLARLTGHLSLASAAHLYLGRPQDFIGGEPTPGRYCHGGFG
VLLSRMLLQQLRPHLEGCRNDIVSARPDEWLGRCIL DATGVGCTGDHEGVHYSHLELSPGEP
VQEGDPHFRSALT AHPVRDPVHMYQLHKAFARAELE RTYQEIQELQWEIQNTSHLAVDGDRA
AAWPVGI PAPS RPASRFEVLRWDYFTEQHAFSCADGSPRCPLRGADRADVADVLGTALEELN
RRYHPALRLQKQQLVNGYRRFDPARGMEYTL DLQLEALTPQGRRPLTRRVQLLRPLSRVEI
LPVPYVTEASRLTVLLPLAAAERDLAPGFLEAFATAALEPGDAAAALTLLLLYEPRQAQRVA
HADVFAPVKAHVAELERRFPGARVPWLSVQTAAPSPLRLMDLLSKKHPLDTLFLLAGPDTVL
TPDFLNRCRMHAISGWQAFFPMHFQAFHPGVAPPQGGPPPELGRDTGRFDRQAASEACFYNS
DYVAARGRLAAASEQEEELLES LDVYELFLHFSSLHVLRAVEPALLQRYRAQTCSARLSEDL
YHRCLQSVLEGLGSRTQLAMLLFEQE QGNST

FIGURE 234

GCTCTGGCCGGCCCCGGCGATTGGTCACCGCCCGCTAGGGGACAGCCCTGGCCTCCTCTGAT
TGGCAAGCGCTGGCCACCTCCCCACACCCCTTGCGAACGCTCCCCTAGTGGAGAAAAGGAGT
AGCTATTAGCCAATTCGGCAGGGCCCGCTTTTTAGAAAGCTTGATTTCTTTGAAGATGAAAG
ACTAGCGGAAGCTCTGCCTCTTTCCCCAGTGGGCGAGGGAACTCGGGGCGATTGGCTGGGAA
CTGTATCCACCCAAATGTCACCGATTTCTTCCTATGCAGGAAATGAGCAGACCCATCAATAA
GAAATTTCTCAGCCTGGCCGAAAATGGTTGGCCCCACGAAGCCACGACAACTGGAGGCAAAG
AGGGTTGCTCAACGCCCCGCCTCATTGGAAAACCAAATCAGATCTGGGACCTATATAGCGTG
GCGGAGGCGGGGCGATGATTGTCGCGCTCGCACCCACTGCAGCTGCGCACAGTCGCATTTCT
TTCCCCGCCCTGAGACCCTGCAGCACCATCTGT**CATG**GCGGCTGGGCTGTTTGGTTTGAGC
GCTCGCCGTCTTTTGGCGGCAGCGGCGACGCGAGGGCTCCCGGCCGCCCGCGTCCGCTGGGA
ATCTAGCTTCTCCAGGACTGTGGTCGCCCCGTCCGCTGTGGCGGGAAAGCGGCCCCCAGAAC
CGACCACACCGTGGCAAGAGGACCCAGAACCCGAGGACGAAAACCTTGATGAGAAGAACCCA
GACTCCCATGGTTATGACAAGGACCCCGTTTTGGACGTCTGGAACATGCGACTTGTCTTCTT
CTTTGGCGTCTCCATCATCCTGGTCCTTGGCAGCACCTTTGTGGCCTATCTGCCTGACTACA
GGATGAAAGAGTGGTCCC GCCGGAAGCTGAGAGGCTTGTGAAATACCGAGAGGCCAATGGC
CTTCCCATCATGGAATCCAACCTGCTTCGACCCCAGCAAGATCCAGCTGCCAGAGGATGAG**TG**
ACCAGTTGCTAAGTGGGGCTCAAGAAGCACCGCCTTCCCCACCCCCTGCCTGCCATTCTGAC
CTCTTCTCAGAGCACCTAATTAAAGGGGCTGAAAGTCTGAA

235/330

FIGURE 235

MAAGLFGLSARLLAAAATRGLPAARVRWESSFSRTVVAPSAVAGKRPEPTTPWQEDPEPE
DENLYEKNPD SHGYDKDPVLDVWNMRLVFFFGVSIILVLGSTFVAYLPDYRMKEWSRREAER
LVKYREANGLPIMESNCFDPSKIQLPED

235/330

FIGURE 236

GGCGGCTGGGCTGTTTGGTTTGAGCGCTCGCCGTCTTTTGGCGGCAGCGGCGACGCGAGGGC
TCCCGGCCGCCCCGCGTCCGCTGGGAATCTAGCTTCTCCAGGACTGTGGTCGCCCCGTCCGCT
GTGGCGGGAAAGCGGCCCCCAGAACCGACCACACCGTGGCAAGAGGACCCAGAACCCGAGGA
CGAAAAC TTGTATGAGAAGAACCCAGACTCCCATGGTTATGACAAGGACCCCGTTTTGGACG
TCTGGAACATGCGACTTGTCTTCTTCTTTGGCGTCTCCATCATCCTGGTCCTTGGCAGCACC
TTTGTGGCCTATCTGCCTGACTACAGGATGAAAGAGTGGTCCCGCCGCGAAGCTGAGAGGCT
TGTGAAATACCGAGAGGCCAATGGCCTTCCCATCATGGAATCCAAC TGCTTCGACCCCAGCA
AGATCCAG

0123456789101112131415161718192021222324252627282930313233343536373839404142434445464748495051525354555657585960616263646566676869707172737475767778798081828384858687888990919293949596979899

FIGURE 238

MPLALLVLLLLGPGGWCLAEP RDSLREELVITPLPSGDVAATFQFRTRWDSELQREGVSHY
RLFPKALGQLISKYSLRELHLSFTQGFWRTRYWGPPFLQAPSGAELWVWFQDTVTDVDKSWK
ELSNVLSGIFCASLNFIDSTNTVTPTASF KPLGLANDTDHYFLRYAVLPREVVCTENLTPWK
KLLPCSSKAGLSVLLKADRLFHTSYHSQAVHIRPVCARNARCTSI SWELRQTL SVVFDAFITG
QGKKDWSLFRMFSRTLTEPCPLASESRVYVDITTYNQDNETLEVHPPPTTTYQDVILGTRKT
YAIYDLLDTAMINNSRNLNIQLKWKRPPE NEAPPVPFLHAQRYVSGYGLQKGELSTLLYNTH
PYRAFPVLLLDTPWYLRLYVHTLTITSKGKENKPSYIHYQPAQDRLQPHLLEMLIQLPANS
VTKVSIQFERALLKWTEYTPDPNHGFYVSPSVLSALVPSMVAAPVDWEESPLFNSLFPVSD
GSNYFVRLYTEPLL VNLPTPDFSMPYNVICLTCTVVAVCYGSFYNLLTRTFHIEEPRTGGLA
KRLANLIRRARGVPPL

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 239

CAACATGGGGTCCAGCAGCTTCTTGGTCCTCATGGTGTCTCTCGTTCTTGTGACCCTGGTGG
CTGTGGAAGGAGTTAAAGAGGGTATAGAGAAAGCAGGGGTTTGCCCAGCTGACAACGTACGC
TGCTTCAAGTCCGATCCTCCCCAGTGTACACAGACCAGGACTGTCTGGGGGAAAGGAAGTG
TTGTTACCTGCACTGTGGCTTCAAGTGTGTGATTCCTGTGAAGGAACTGGAAGAAGGAGGAA
ACAAGGATGAAGATGTGTCAAGGCCATACCCTGAGCCAGGATGGGAGGCCAAGTGTCCAGGC
TCCTCCTCTACCAGGTGTCCTCAGAAATGAATGCTGGGTCTTTCTACCTCTGGGGGTCCTC
TCACTTGGCACCTGCCCCCTGAGGGTCCTGAGACTTGGAATATGGAAGAAGCAATACCCAACC
CCACCAAAGAAAACCTGAGCTTGAAGTCCTTTTCCCCAAAAGAGGGAAGAGTCACAAAAAG
TCCAGACCCCAGGGACGGTACTTTCCCTCTCTACCTGGTGCTCCTCCCTAATGCTCATGAAT
GGACCCCTCATGAATGAAACCAGTGCCCTTATAAGAGACCCCAAAGAGCTGCCTTGCCCTTC
TGCAATGTGTGATCACAGCTAGAAGGCACTGTCAGAGAAGAGAACTGGTCCTCACCAGATG
CTGAATCTGCTGGTGCCTTGATCTTGGACTTCCCAGCCTCTAGAACTGTAAGAAATAAATAT
TTGCTGTTTATAATCCAA

240/330

FIGURE 240

MGSSSFVLVLMVSLVLVTLVAVEGVKEGIEKAGVCPADNVRCFKSDPPQCHTDQDCLGERKCC
YLHCGFKCVIPVKELEEKGNKDEDVSRPYPEPGWEAKCPGSSSTRCPQK

Signal sequence:

amino acids 1-19

N-myristoylation sites:

amino acids 23-29, 27-33, 32-38, 102-108

WAP-type 'four-disulfide core' domain signature:

amino acids 49-63

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

FIGURE 242

MRSCLWRCRHLSQGVQWSLLLAVLVFFLFALPSFIKEPQTKPSRHQRTENIKERSLQSLAKP
 KSQAPTRARRTTIYAEPAPENNALNTQTQPKAHTTGDRGKEANQAPPEEQDKVPHTAQRAAW
 KSPEKEKTMVNTLSPRGQDAGMASGRTEAQSWKSQDTKTTQGNNGQTRKLTASRTVSEKHQG
 KAATTAKTLIPKSQHRMLAPTGA VSTRTRQKGVTTAVIPPKEKKPQATPPPAPFQSPTTQRN
 QRLKAANFKSEPRWDFEEKYSFEIGGLQTTCPDSVKIKASKSLWLQKLFLPNLTLFLDSRHF
 NQSEWDRLEHFAPPFGEFMEFNYSLVQKVVTFRFPVPQQQLLLASLPAGSLRCITCAVVGNGG
 ILNNSHMGQEIDSHDYVFERLSGALIKGYEQDVGTRTSFYGFATAFSLTQSLILGNRGFKNVP
 LGKDVRYLHFLEGTRDYEWLEALLMNQTVMSKNLFWFRHRPQEAFFREALHMDRYLLLHPDFL
 RYMKNRFLRSKTL DGAHWRIYRPTTGALLLLTALQLCDQVSAYGFITEGHERFSDHYDTSW
 KRLIFYINHDFKLEREVWKRLHDEGIIRLYQRPGPGTAKAKN

Cytoplasmic Domain:

amino acids 1-10

Type II Transmembrane Domain:

amino acids 11-35

Lumenal catalytic Domain:

amino acids 36-600

Ribonucleotide Reductase small subunit Signature:

amino acids 481-496

N-glycosylation Sites:

amino acids 300-303, 311-314, 331-334, 375-378, 460-463

FIGURE 243

CGATGCGCGGACCCGGGCACCCCCTCCTCCTGGGGCTGCTGCTGGTGCTGGGGCCTTCGCCG
GAGCAGCGAGTGGAAATTGTTCTCGAGATCTGAGGATGAAGGACAAGTTTCTAAACACCT
TACAGGCCCTCTTTATTTTAGTCCAAAGTGCAGCAAACACTTCCATAGACTTTATCACAACA
CCAGAGACTGCACCATTCCTGCATACTATAAAAGATGCGCCAGGCTTCTTACCCGGCTGGCT
GTCAGTCCAGTGTGCATGGAGGATAAGTGAGCAGACCGTACAGGAGCAGCACACCAGGAGCC
ATGAGAAGTGCCTTGGAACCAACAGGGAAACAGAACTATCTTTATACACATCCCCTCATGG
ACAAGAGATTTATTTTGGCAGACAGACTCTTCCATAAGTCCTTTGAGTTTTGTATGTTGTTG
ACAGTTTGCAGATATATATTGATAAATCAGTGTACTTGACAGTGTTATCTGTCACTTATTT

CGCGGACCCGGGCACCCCCTCCTCCTGGGGCTGCTGCTGGTGCTGGGGCCTTCGCCG
GAGCAGCGAGTGGAAATTGTTCTCGAGATCTGAGGATGAAGGACAAGTTTCTAAACACCT
TACAGGCCCTCTTTATTTTAGTCCAAAGTGCAGCAAACACTTCCATAGACTTTATCACAACA
CCAGAGACTGCACCATTCCTGCATACTATAAAAGATGCGCCAGGCTTCTTACCCGGCTGGCT
GTCAGTCCAGTGTGCATGGAGGATAAGTGA

FIGURE 245

GGGCTGGGCCCCGCCGCAGCTCCAGCTGGCCGGCTTGGTCCTGCGGTCCCTTCTCTGGGAGG
CCCGACCCCGGCCGCGCCAGCCCCACCATGCCACCCGCGGGGCTCCGCCGGGCCGCGCCG
CTCACCGCAATCGCTCTGTTGGTGCTGGGGGCTCCCCTGGTGCTGGCCGGCGAGGACTGCCT
GTGGTACCTGGACCGGAATGGCTCCTGGCATCCGGGGTTTAACTGCGAGTTCTTCACCTTCT
GCTGCGGGACCTGCTACCATCGGTACTGCTGCAGGGACCTGACCTTGCTTATCACCGAGAGG
CAGCAGAAGCACTGCCTGGCCTTCAGCCCCAAGACCATAGCAGGCATCGCCTCAGCTGTGAT
CCTCTTTGTTGCTGTGGTTGCCACCACCATCTGCTGCTTCCTCTGTTCTGTTGCTACCTGT
ACCGCCGGCGCCAGCAGCTCCAGAGCCCATTGAAGGCCAGGAGATTCCAATGACAGGCATC
CCAGTGCAGCCAGTATACCCATACCCCAAGGACCCCAAGCTGGCCCTGCACCCCAAGCC
TGGCTTCATGTACCCACCTAGTGGTCCTGCTCCCCAATATCCACTCTACCCAGCTGGGCCCC
CAGTCTACAACCCTGCAGCTCCTCCTCCCTATATGCCACCACAGCCCTCTTACCCGGGAGCC
TGAGGAACCAGCCATGTCTCTGCTGCCCCCTTCAGTGATGCCAACCTTGGGAGATGCCCTCAT
CCTGTACCTGCATCTGGTCCTGGGGGTGGCAGGAGTCCTCCAGCCACCAGGCCCCAGACCAA
GCCAAGCCCTGGGCCCTACTGGGGACAGAGCCCCAGGGAAGTGGAACAGGAGCTGAACTAGA
ACTATGAGGGGTGGGGGGAGGGCTTGGAATTATGGGCTATTTTTACTGGGGGCAAGGGAGG
GAGATGACAGCCTGGGTCACAGTGCCTGTTTTCAAATAGTCCCTCTGCTCCCAAGATCCCAG
CCAGGAAGGCTGGGGCCCTACTGTTTGTCCCCTCTGGGCTGGGGTGGGGGGAGGGAGGAGGT
TCCGTCAGCAGCTGGCAGTAGCCCTCCTCTCTGGCTGCCCCACTGGCCACATCTCTGGCCTG
CTAGATTAAAGCTGTAAAGACAAAA

FIGURE 246

MPPAGLRRAAPLTAIALLLVLGAPLVLAGEDCLWYLDNRNGSWHPGFNCEFFTFC CGTCYHRYC
CRDLTLLITERQQKHCLAFSPKTIAGIASAVILFVAVVATTICCF LCSCCYLYRRRQQLQSP
FEGQEIPMTGIPVQPVYPYPQDPKAGPAPPQPGFMYP PPSGPAPQYPLYPAGPPVYNPAAPP
YMPPQPSYPGA

Transmembrane Domains:

amino acids 10-28, 85-110

N-glycosylation Site:

amino acids 38-41

N-myristoylation Sites:

amino acids 5-10, 88-93

FIGURE 247

GGGGGAGCTAGGCCGGCGGCAGTGGTGGTGGCGGCGGCGCAAGGGTGAGGGCGGCCCCAGAA
 CCCCAGGTAGGTAGAGCAAGAAG**ATGGT**GTTTTCTGCCCTCAAATGGTCCCTTGCAACCATG
 TCATTTCTACTTTCTCCTCACTGTTGGCTCTCTTAAGTGTGTCCACTCCTTCATGGTGTGAGAG
 CACTGAAGCATCTCCAAAACGTAGTGATGGGACACCATTTTCTTGGAATAAAATACGACTTC
 CTGAGTACGTCATCCAGTTTCATTATGATCTCTTGATCCATGCAAACCTTACCACGCTGACC
 TTCTGGGGAACACGAAAGTAGAAATCACAGCCAGTCAGCCCACCAGCACCATCATCTGCA
 TAGTCACCACCTGCAGATATCTAGGGCCACCTCAGGAAGGGAGCTGGAGAGAGGCTATCGG
 AAGAACCCCTGCAGGTCTGGAACACCCCCCTCAGGAGCAAATTGCACTGCTGGCTCCCGAG
 CCCCTCCTTGTGCGGCTCCCGTACACAGTTGTCACTTCACTATGCTGGCAATCTTTCGGAGAC
 TTTCCACGGATTTTACAAAAGCACCTACAGAACCAAGGAAGGGGAACTGAGGATACTAGCAT
 CAACACAATTTGAACCCACTGCAGCTAGAATGGCCTTTCCCTGCTTTGATGAACCTGCCTTC
 AAAGCAAGTTTCTCAATCAAATTAGAAGAGAGCCAAGGCACCTAGCCATCTCCAATATGCC
 ATTGGTGAAATCTGTGACTGTTGCTGAAGGACTCATAGAAGACCATTTTGATGTCACTGTGA
 AGATGAGCACCTATCTGGTGGCCTTCATCATTTTCAGATTTTGAGTCTGTGAGCAAGATAACC
 AAGAGTGGAGTCAAGTTTTCTGTTTATGCTGTGCCAGACAAGATAAATCAAGCAGATTATGC
 ACTGGATGCTGCGGTGACTCTTCTAGAATTTTATGAGGATTATTTTCAGCATACCGTATCCCC
 TACCCAAACAAGATCTTGCTGCTATTTCCGACTTTTCAGTCTGGTGTATGGAAAACCTGGGGA
 CTGACAACATATAGAGAATCTGCTCTGTTGTTGATGCAAGAAAGTCTTCTGCATCAAGTAA
 GCTTGGCATCACAGTGACTGTGGCCCATGAACTGGCCACCAGTGGTTTGGGAACCTGGTCA
 CTATGGAATGGTGGAAATGATCTTTGGCTAAATGAAGGATTTGCCAAATTTATGGAGTTTGTG
 TCTGTGAGTGTGACCCATCCTGAACTGAAAGTTGGAGATTATTTCTTTGGCAAATGTTTTGA
 CGCAATGGAGGTAGATGCTTTAAATTCCTCACACCTGTGTCTACACCTGTGGAAAATCCTG
 CTCAGATCCGGGAGATGTTTGATGATGTTTCTTATGATAAGGGAGCTTGTATTCTGAATATG
 CTAAGGGAGTATCTTAGCGCTGACGCATTTAAAGTGGTATTGTACAGTATCTCCAGAAGCA
 TAGCTATAAAAATACAAAAACGAGGACCTGTGGGATAGTATGGCAAGTATTTGCCCTACAG
 ATGGTGTAAAAGGGATGGATGGCTTTTGCTCTAGAAGTCAACATTCATCTTCATCCTCACAT
 TGGCATCAGGAAGGGGTGGATGTGAAAACCATGATGAACACTTGGACACTGCAGAGGGGTTT
 TCCCCTAATAACCATCACAGTGAGGGGGAGGAATGTACACATGAAGCAAGAGCACTACATGA
 AGGGCTCTGACGGCGCCCCGGACACTGGGTACCTGTGGCATGTTCCATTGACATTCATCACC
 AGCAAATCCAACATGGTCCATCGATTTTTGCTAAAAACAAAAACAGATGTGCTCATCCTCCC
 AGAAGAGGTGGAATGGATCAAATTTAATGTGGGCATGAATGGCTATTACATTGTGCATTACG
 AGGATGATGGATGGGACTCTTTGACTGGCCTTTTAAAGGAACACACACAGCAGTCAGCAGT
 AATGATCGGGCAAGTCTCATTAACAATGCATTTTCAGCTCGTCAGCATTGGGAAGCTGTCCAT
 TGAAAAGGCCCTTGGAATTTATCCCTGTACTTGAAACATGAACTGAAATTATGCCCGTGTTC
 AAGGTTTGAATGAGCTGATTCCTATGTATAAGTTAATGGAGAAAAGAGATATGAATGAAGTG
 GAAACTCAATTCAAGGCCTTCCTCATCAGGCTGCTAAGGGACCTCATTGATAAGCAGACATG
 GACAGACGAGGGCTCAGTCTCAGAGCAAATGCTGCGGAGTGAACACTACTACTCCTCGCCTGTG
 TGCACAACTATCAGCCGTGCGTACAGAGGGCAGAAGGCTATTTTCAGAAAGTGGAAGGAATCC
 AATGGAACTTGAGCCTGCCTGTGACGTGACCTTGGCAGTGTTTGCTGTGGGGGCCAGAG
 CACAGAAGGCTGGGATTTTCTTTATAGTAAATATCAGTTTTCTTTGTCCAGTACTGAGAAAA
 GCCAAATTGAATTTGCCCTCTGCAGAACCCAAAATAAGGAAAAGCTTCAATGGCTACTAGAT
 GAAAGCTTTAAGGGAGATAAAATAAAAACTCAGGAGTTTCCACAAATTTCTTACACTCATTGG
 CAGGAACCCAGTAGGATACCCACTGGCCTGGCAATTTCTGAGGAAAACTGGAACAACTTG
 TACAAAAGTTTGAACCTGGCTCATCTTCCATAGCCACATGGTAATGGGTACAACAAATCAA
 TTCTCCACAAGAACACGGCTTGAAGAGGTAAAAGGATTCTTCAGCTCTTTGAAAGAAAATGG
 TTCTCAGCTCCGTTGTGTCCAACAGACAATTGAAACCATTGAAGAAAACATCGGTTGGATGG
 ATAAGAATTTTGATAAAATCAGAGTGTGGCTGCAAAGTGAAAAGCTTGAACGTATG**TAAAAA**
 TTCCCTCCCTTGCCCGGTTCTGTTATCTCTAATCACCAACATTTTGTGAGTGTATTTTCAA
 ACTAGAGATGGCTGTTTTGGCTCCAACCTGGAGATACTTTTTTCCCTTCAACTCATTTTTTGA
 CTATCCCTGTGAAAAGAATAGCTGTTAGTTTTTCATGAATGGGCTTTTTTCATGAATGGGCTA
 TCGCTACCATGTGTTTTGTTTCATCACAGGTGTTGCCCTGCAACGTAAACCCAAGTGTTGGGT
 TCCCTGCCACAGAAGAATAAAGTACCTTATTCTCTCAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 248

MVFLPLKWSLATMSFLLSSLLALLTVSTPSWCQSTEASPKRSDGTPFPWNKIRLPEYVIPVH
 YDLLIHANLTTLTFWGTTKVEITASQPTSTIILHSHHLQISRATLRKGAGERLSEEPLQVLE
 HPPQEQIALLAPEPLLVGLPYTVVIHYAGNLSETFHGFYKSTYRTKEGELRILASTQFEPTA
 ARMAFPCFDEPAFKASFSEIKIRREPRHLAISNMPLVKSVTVAEGLIEDHFDVTVKMSTYLVA
 FIISDFESVSKITKSGVKVSVYAVDPKINQADYALDAAVTLLEFYEDYFSIPYPLPKQDLAA
 IPDFQSGAMENWGLTTYRESALLFDAEKSSASSKLGITVTVAHELAHQWFGNLVTMEWWNDL
 WLNEGFAKFMEFVSVSVTHPELKVG DYFFGKCFDAMEVDALNSSHPVSTPVENPAQIREMFD
 DVSYDKGACILNMLREYLSADAFKSGIVQYLQKHSYKNTKNEDLWDSMASICPTDGVKGMDG
 FCSRSQHSSSSSHWHQEGVDVKTMMNTWTLQRGFPLITITVRGRNVHMKQEHYMKGSDGAPD
 TGYLWHVPLTFITSKSNMVHRFLKTKTDVLILPEEVEWIKFNVGMNGYYIVHYEDDGWDSL
 TGLLKGTHTAVSSNDRASLINNAFQLVSIKLSIEKALDLSLYLKHETEIMPVFQGLNELIP
 MYKLMEKRDMEVETQFKAFILRLRLDLIDKQWTDEGSVSEQMLRSELLLLACVHNYQPCV
 QRAEGYFRKWKESNGNLSLPVDVTLAVFAVGAQSTEGWDFLYSKYQFSLSSTEKSQIEFALC
 RTQNKEKLQWLLDESFKGDKIKTQEFQILTILGRNPVGYPLAWQFLRKNWNKLVQKFELGS
 SSIAHMVMGTTNQFSTRTRLEEVKGFFSSSLKENGSQLRCVQQTETIETIENIGWMDKNFDKIR
 VWLQSEKLERM

Signal peptide:

amino acids 1-34

N-glycosylation sites:

amino acids 70-74, 154-158, 414-418, 760-764, 901-905

Neutral zinc metallopeptidases, zinc-binding region signature:

amino acids 350-360

FIGURE 249

CAGCCACAGACGGGTC**ATG**AGCGCGGTATTACTGCTGGCCCTCCTGGGGTTCATCCTCCCAC
TGCCAGGAGTGCAGGCGCTGCTCTGCCAGTTTGGGACAGTTCAGCATGTGTGGAAGGTGTCC
GACCTACCCCGGCAATGGACCCCTAAGAACACCAGCTGCGACAGCGGCTTGGGGTGCCAGGA
CACGTTGATGCTCATTTGAGAGCGGACCCCAAGTGAGCCTGGTGCTCTCCAAGGGCTGCACGG
AGGCCAAGGACCAGGAGCCCCGCGTCACTGAGCACCGGATGGGCCCCGGCCTCTCCCTGATC
TCCTACACCTTCGTGTGCCGCCAGGAGGACTTCTGCAACAACCTCGTTAACTCCCTCCCGCT
TTGGGCCCCACAGCCCCCAGCAGACCCAGGATCCTTGAGGTGCCAGTCTGCTTGTCTATGG
AAGGCTGTCTGGAGGGGACAACAGAAGAGATCTGCCCCAAGGGGACCACACACTGTTATGAT
GGCCTCCTCAGGCTCAGGGGAGGAGGCATCTTCTCCAATCTGAGAGTCCAGGGATGCATGCC
CCAGCCAGGTTGCAACCTGCTCAATGGGACACAGGAAATTGGGCCCCGTGGGTATGACTGAGA
ACTGCAATAGGAAAGATTTTCTGACCTGTCATCGGGGGACCACCATTATGACACACGGAAAC
TTGGCTCAAGAACCCACTGATTGGACCACATCGAATACCGAGATGTGCGAGGTGGGGCAGGT
GTGTCAGGAGACGCTGCTGCTCATAGATGTAGGACTCACATCAACCCTGGTGGGGACAAAAG
GCTGCAGCACTGTTGGGGCTCAAAATTCCCAGAAGACCACCATCCACTCAGCCCCCTCCTGGG
GTGCTTGTGGCCTCCTATACCCACTTCTGCTCCTCGGACCTGTGCAATAGTGCCAGCAGCAG
CAGCGTTCTGCTGAACTCCCTCCCTCCTCAAGCTGCCCCGTGCCAGGAGACCGGCAGTGTC
CTACCTGTGTGCAGCCCCCTTGGAACCTGTTCAAGTGGCTCCCCCGAATGACCTGCCCCAGG
GGCGCCACTCATTGTTATGATGGGTACATTATCTCTCAGGAGGTGGGCTGTCCACCAAAT
GAGCATTACAGGGCTGCGTGGCCCCAACCTTCCAGCTTCTTGTGAACCACACCAGACAAATCG
GGATCTTCTCTGCGCGTGAGAAGCGTGATGTGCAGCCTCCTGCCTCTCAGCATGAGGGAGGT
GGGGCTGAGGGCCTGGAGTCTCTCACTTGGGGGGTGGGGCTGGCACTGGCCCCAGCGCTGTG
GTGGGGAGTGGTTTGCCCTTCCTGC**TAA**CTCTATTACCCCCACGATTCTTCACCGCTGCTGA
CCACCCACACTCAACCTCCCTCTGACCTCATAACCTAATGGCCTTGGACACCAGATTCTTTC
CCATTCTGTCCATGAATCATCTTCCCCACACACAATCATTCATATCTACTCACCTAACAGCA
AACTGGGGAGAGCCTGGAGCATCCGGACTTGCCCTATGGGAGAGGGGACGCTGGAGGAGTG
GCTGCATGTATCTGATAATACAGACCCTGTCCTTTCA

FIGURE 250

MSAVLLLALLGFILPLPGVQALLCQFGTVQHVWKVSDLPRQWTPKNTSCDSGLGCQDTLMLI
ESGPQVSLVLSKGCTEAKDQEPRVTEHRMGPGLSLISYTFVCRQEDFCNNLVNSLPLWAPQP
PADPGSLRCPVCLSMEGCLEGTTEEICPKGTHCYDGLLRRLRGGGIFSNLRVQGCMPQPGCN
LLNGTQEIGPVGMTENCNRKDFLTCHRGTTIMTHGNLAQEPTDWTTSNTEMCEVGQVCQETL
LLIDVGLTSTLVGTKGCSTVGAQNSQKTTIHSAPPGVLVASYTHFCSSDLCNSASSSSVLLN
SLPPQAAPVPGDRQCPTCVQPLGTCSSGSPRMTCPRGATHCYDGYIHLSGGGLSTKMSIQGC
VAQPSSFLLNHTRQIGIFSAREKRDVQPPASQHEGGGAEGLESLTWGVGLALAPALWWGVVC
PSC

250/330

FIGURE 251

GCGACGGGCAGGACGCCCCGTTTCGCCTAGCGCGTGCTCAGGAGTTGGTGTCCCTGCCTGCGCT
CAGG**ATG**AGGGGGAATCTGGCCCTGGTGGGCGTTCTAATCAGCCTGGCCTTCCTGTCACTGCTG
CCATCTGGACATCCTCAGCCGGCTGGCGATGACGCCTGCTCTGTGCAGATCCTCGTCCCTGG
CCTCAAAGGGGATGCGGGAGAGAAGGGAGACAAAGGCGCCCCGGACGGCCTGGAAGAGTCG
GCCCCACGGGAGAAAAAGGAGACATGGGGGACAAAGGACAGAAAGGCAGTGTGGGTGCTCAT
GGAAAAATTGGTCCCATTTGGCTCTAAAGGTGAGAAAGGAGATTCCGGTGACATAGGACCCCC
TGGTCCTAATGGAGAACCAGGCCTCCCATGTGAGTGCAGCCAGCTGCGCAAGGCCATCGGGG
AGATGGACAACCAGGTCTCTCAGCTGACCAGCGAGCTCAAGTTCATCAAGAATGCTGTGCC
GGTGTGCGCGAGACGGAGAGCAAGATCTACCTGCTGGTGAAGGAGGAGAAGCGCTACGCGGA
CGCCCAGCTGTCTGCCAGGGCCGCGGGGGCACGCTGAGCATGCCCAAGGACGAGGCTGCCA
ATGGCCTGATGGCCGCATACCTGGCGCAAGCCGGCCTGGCCCCGTGTCTTCATCGGCATCAAC
GACCTGGAGAAGGAGGGCGCCTTCGTGTACTCTGACCACTCCCCCATGCGGACCTTCAACAA
GTGGCGCAGCGGTGAGCCCAACAATGCCTACGACGAGGAGGACTGCGTGGAGATGGTGGCCT
CGGGCGGCTGGAACGACGTGGCCTGCCACACCACCATGTACTTCATGTGTGAGTTTGACAAG
GAGAACATG**TGA**GCCTCAGGCTGGGGCTGCCCATTGGGGGCCCCACATGTCCCTGCAGGGTT
GGCAGGGACAGAGCCCAGACCATGGTGCCAGCCAGGGAGCTGTCCCTCTGTGAAGGGTGGAG
GCTCACTGAGTAGAGGGCTGTTGTCTAAACTGAGAAAATGGCCTATGCTTAAGAGGAAAATG
AAAGTGTTCCCTGGGGTGCTGTCTCTGAAGAAGCAGAGTTTCATTACCTGTATTGTAGCCCCA
ATGTCATTATGTAATTATTACCCAGAATTGCTCTTCCATAAAGCTTGTGCCTTTGTCCAAGC
TATACAATAAAATCTTTAAGTAGTGCAGTAGTTAAGTCCAAAAAAAAAAAAAAAAAAAAA

FIGURE 252

MRGNLALVGVLI SLAFLSLLPSGHPQPAGDDACSVQILVPGLKGDAGEKGDKGAPGRPGRVG
PTGEKGDMDKGQKGSVGRHGKIGPIGSKGEKGDSDIGPPGPNGEPGLPCECSQLRKAIGE
MDNQVSQLTSELKFIKNAVAGVRETESKIYLLVKEEKRYADAQLSCQGRGGTLSMPKDEAAN
GLMAAYLAQAGLARVFIGINDLEKEGAFVYSDHSPMRTFNKWRSGEPNNAYDEEDCVEMVAS
GGWNDVACHTTMYFMCEFDKENM

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 253

AGTGACTGCAGCCTTCCTAGATCCCCTCCACTCGGTTTCTCTCTTTGCAGGAGCACCGGCAG
CACCAGTGTGTGAGGGGAGCAGGCAGCGGTCCTAGCCAGTTCCTTGATCCTGCCAGACCACC
CAGCCCCCGGCACAGAGCTGCTCCACAGGCACC**ATG**AGGATCATGCTGCTATTCACAGCCAT
CCTGGCCTTCAGCCTAGCTCAGAGCTTTGGGGCTGTCTGTAAGGAGCCACAGGAGGAGGTGG
TTCCTGGCGGGGGCCGCAGCAAGAGGGATCCAGATCTCTACCAGCTGCTCCAGAGACTCTTC
AAAAGCCACTCATCTCTGGAGGGATTGCTCAAAGCCCTGAGCCAGGCTAGCACAGATCCTAA
GGAATCAACATCTCCCGAGAAACGTGACATGCATGACTTCTTTGTGGGACTTATGGGCAAGA
GGAGCGTCCAGCCAGAGGGAAAGACAGGACCTTTCTTACCTTCAGTGAGGGTTCCTCGGCCC
CTTCATCCCAATCAGCTTGGATCCACAGGAAAGTCTTCCCTGGGAACAGAGGAGCAGAGACC
TTT**TAA**GACTCTCCTACGGATGTGAATCAAGAGAACGTCCCCAGCTTTGGCATCCTCAAGT
ATCCCCCGAGAGCAGAATAGGTACTCCACTTCCGGACTCCTGGACTGCATTAGGAAGACCTC
TTTCCCTGTCCCAATCCCCAGGTGCGCACGCTCCTGTTACCCTTTCTCTTCCCTGTTCTTGT
AACATTCTTGTGCTTTGACTCCTTCTCCATCTTTTCTACCTGACCCTGGTGTGGAAACTGCA
TAGTGAATATCCCCAACCCCAATGGGCATTGACTGTAGAATACCCTAGAGTTCCTGTAGTGT
CCTACATTAAAAATATAATGTCTCTCTCTATTCCCTCAACAATAAAGGATTTTTGCATATGAA
AA

FIGURE 254

MRIMLLFTAILAFSLAQSFQAVCKEPQEEVVPGGGRSKRDPDLYQLLQRLFKSHSSLEGLLK
ALSQASTDPKESTSPEKRDMHDFVGLMGKRSVQPEGKTGPFLLPSVRVPRPLHPNQLGSTGK
SSLGTEEQRPL

Important features:

Signal peptide:

amino acids 1-18

Tyrosine kinase phosphorylation site.

amino acids 36-45

N-myristoylation site.

amino acids 33-39, 59-65

Amidation site.

amino acids 90-94

Leucine zipper pattern.

amino acids 43-65

Tachykinin family signature.

amino acids 86-92

FIGURE 255

GGGCGTCTCCGGCTGCTCCTATTGAGCTGTCTGCTCGCTGTGCCCCGCTGTGCCTGCTGTGCC
CGCGCTGTGCGCGCTGCTACCGCGTCTGCTGGACGCGGGAGACGCCAGCGAGCTGGTGATTG
GAGCCCTGCGGAGAGCTCAAGCGCCCAGCTCTGCCCCAGGAGCCCAGGCTGCCCCGTGAGTC
CCATAGTTGCTGCAGGAGTGGAGCCATGAGCTGCGTCCTGGGTGGTGTCATCCCCCTTGGGGC
TGCTGTTCCCTGGTCTGCGGATCCCAAGGCTACCTCCTGCCCAACGTCACTCTCTTAGAGGAG
CTGCTCAGCAAATACCAGCACAAAGAGTCTCACTCCCGGGTCCGCAGAGCCATCCCCAGGGA
GGACAAGGAGGAGATCCTCATGCTGCACAACAAGCTTCGGGGCCAGGTGCAGCCTCAGGCCT
CCAACATGGAGTACATGGTGAGCGCCGGCTCCGGCCGCAGAGGCTGGCACCGGGGGTGGGGC
CTGGGCCACCAGCCTGCTCTGTTCCCCAGCCAGCTCTGTTCCCCAGCCAGTGCGTGTGATGG
CTGGCTCAGGGTCTCCTCTGGCAGGGGAGGATCCCGGCTCTGTTCTGTTTTGTTTGTGTTGTT
TTGAGACAGGGTCTCACTCTGCCACTGACGCTGGAGTGCAATGGCACAATCGTCATGCCCTG
AAACCTTAGACTCCCGGGGTAAAGCGATCCTGCTTCAGCCTCCCAAGTAGCTGGAACCTACAG
GCATGCACCATGGTGCCAGCTAGATTTTAAATATTTTGTGGAGATGGGGGTCTTGCTACGT
TGCCCAGGCTGGTCTTGAACCTCCTAGGCTCAAGCAATCCTCCTGCCTCAGCCTCTCAAAGTG
CTAGGATTATAGGCATGAGTCACCCTGTCTGGCTCTGGCTCTGTTCTTAACATTCTGCCAAA
ACAACACACGTGGGTTCCTGTGCAGAGCCTGCCTCGTTGCCTTCATGTCACTCTTGGTAGC
TCCACTGGGAACACAGCTCTCAGCCTTTCCACCTGGAGGCAGAGTGGGGAGGGGGCCCAGGG
CTGGGCTTTGCTGATGCTGATCTCAGCTGTGCCACACGCTAGCTGCACCACCCTGACTTCTC
CTTAGCCCGTGTGAGCCTCACTTTCCACTTGGAGAGTCCTTCCTCGCGTGGTTGCCATGACT
GTGAGATAAGTCGAGGCTGTGAAGGGCCCGGCACAGACTGACCTGCCTCCCCAACCCCTAGG
CTTTGCTAACCGGGAAAGGAGCTAACGGTGACAGAAGACAGCCAAGGTCAACCCTCCCGGGT
GATTGTGATGGGTGTTCCAGGTGTGGTTGGGCGATGCTGCTACTTGACCCCAAGCTCCAGTG
TGGAAACTTCCTTCCTGGCTGGTTTTCCAGAACTACAGAGGAATGGACCACAGTCTTCCAGG
GTCCCTCCTCGTCCACCAACCGGGAGCCTCCACCTTGGCCATCCGTCAGCTATGAATGGCTT
TTTAAACAAACCCACGTCCCAGCCTGGGTAACATGGTAAAGCCCCGTCTCTACAAAAAATC
CAAGTTAGCCGGGCATGGTGGTGCGCACCTGTAGTCCCAGCTGCAGTGGGACTGAGGTGGAG
GTGGAGGTGGGGGTGGGAGCTGAGGAAGGAGGATCGCTTGAGCCTGGGAAGTCGAGGCTGC
AGTGAGCTGAGATTGCACCACTGCACTCCAGCCTGGGTGACAGAGCAAGACCCTGTCTCAAAAA

FIGURE 256

MSCVLGGV IPLGLLFLVCGSQGYLLPNVTLLLEELLSKYQHNEHSRVRRAIPREDKEEILML
HNKLRGQVQPQASNMEYMVSAGSGRRGWHRGWGLGHQPALFPSQLCSPASACDGLRVSSGR
GGSRLCSVLFVCFETGSHSATDAGVQWHNRHALKP

Important features:

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 27-31, 41-45

N-myristoylation site.

amino acids 126-132, 140-146

Amidation site.

amino acids 85-89

256/330

FIGURE 257

AAGGAGAGGCCACCGGGACTTCAGTGTCTCCTCCATCCCAGGAGCGCAGTGGCCACT**ATGGG**
GTCTGGGCTGCCCCCTTGTCTCCTCTTGACCCTCCTTGGCAGCTCACATGGAACAGGGCCGG
GTATGACTTTGCAACTGAAGCTGAAGGAGTCTTTTCTGACAAATTCCTCCTATGAGTCCAGC
TTCCTGGAATTGCTTGAAAAGCTCTGCCTCCTCCTCCATCTCCCTTCAGGGACCAGCGTCAC
CCTCCACCATGCAAGATCTCAACACCATGTTGTCTGCAACACA**TGA**CAGCCATTGAAGCCTG
TGTCTTCTTGGCCCCGGGCTTTTGGGCCGGGGATGCAGGAGGCAGGCCCCGACCCTGTCTTT
CAGCAGGCCCCCACCCTCCTGAGTGGCAATAAATAAAATTCGGTATGCTG

FIGURE 259

AATTGTATCTGTGTAATGTTAAAACAAACGAAATAAAATAGAAGGAAAACTTTCTGAGTTT
CAAAAACAACAGACTAGTACTCTAAAGAACTCTTTAAAACAATTAAGTGTAGGATTGCAGT
TATGATTGGATATTATTTAATTCTGTTTCTGATGTGGGGTTCCTCCACTGTGTTCTGTGTGC
TATTAATATTTACCATTCGAGAAGCTTCATTCAGTGTGAAAATGAATGCTTAGTGGATCTG
TGCTCTTACGCATATGTTACAAATTATCTGGAGTTCCTAATCAATGCAGAGTCCCCCTCCC
CTCCGATTGTTCTAAATAAATTGAAAGATGTCTGCTGTGGAAAAGGCATGTATTTAAATCTG
TATGATTCTCAACCATCTTTAGTTGGGAAAGGTCCTTGAAAGCCAATGGAAATACTTTTTTTT
TTTTCTTGGCACTAATCAAGTGAGTGTTACCTTTTCACTTAGTAGGATGTGTTGTTACGCTA
GTAAAATAGAAACCTGTGTTTATTCTCAGGTATTTAGAAACAACAGCCATCATTTTATTTT
ATGTGTGTGTTCTTGGCTGTATTCATAAATTATATATTTTGGGCTATCAAATATTACTTCAT
TCAATATAAATAACAATAGTAGAAGTTGTTTACTTAGATATGCTTCTAGTTGCATTTTCTC
AGCCTATGTAAGACTACTTTGTTGTAATAGCCTTTGAAATTTACAGTACTGTCTCTCTACTA
TCTTCAGATTACTTGATTCAAATAAACCAATTATGTTTGTAATTGATATTAATAAAACCAGA
ATAAAAGTTCATATCTACCC

260/330

FIGURE 260

MIGYYLILFLMWGSSTVFCVLLIFTIAEASFVENECLVDLCLLRICYKLSGVPNQCRVPLP
SDCSK

Important features:

Signal peptide:

amino acids 1-29

260/330

FIGURE 262

MMVALRGASALLVLFLAAFLPPPQCTQDPAMVHYIYQRFVLEQGLEKCTQATRAYIQEFQE
FSKNISVMLGRCQTYTSEYKSAVGNLALRVERAQREIDYIQYLREADECIVSEDKTLAEMLL
QEAEEEKKIRTLNASC DNMLMGIKSLKIVKKMMDTHGSWMKDAVYN SPKVYLLIGSRNNTV
WEFANIRAFMEDNTKPAPRKQILTLSWQGTGQVIYKGFLFFHNQATSNEIIKYNLQKRTVED
RMLLPGGVGRALVYQHSPSTYIDLAVDEHGLWAIHSGPGTHSHLVLT KIEPGTLGVEHSWDT
PCRSQDAEASFLLCGVLYVVYSTGGQGPHRITCIYDPLGTISEEDLPNLFFPKRPRSHSMIH
YNPRDKQLYAWNEGNQIIYKLQTKRKLPLK

262/330

FIGURE 264

MELSQMSELMGLSVLLGLLALMATAAVARGWLRAGEERSGRPACQKANGFPPDKSSGSKKQK
QYQRIRKEKPQQHNFTHRLLAAALKSHSGNISCMDFSSNGKYLATCADDRTIRIWSTKDFLQ
REHRSMRANVELDHATLVRFSPDCRAFIVWLANGDTLRVFKMTKREDGGYTFTATPEDFPPK
HKAPVIDIGIANTGKFIMTASSDTTVLIWSLKGQVLSTINTNQMNNTHAAVSPCGRFVASC
FTPDVKVWEVCFGKKGEFQEVVRAFELKGHSAAVHSFAFSNDSRRMASVSKDGTWKLWDTDV
EYKKKQDPYLLKTGRFEEAAGAAPCRLALSPNAQVLALASGSSIHLYNTRRGEKEECFERVH
GECIANLSFDITGRFLASCGDRAVRLFHNTPGHRAMVEEMQGHLKRASNESTRQRLQQQLTQ
AQETLKSLGALKK

Important features:**Signal peptide:**

amino acids 1-25

N-glycosylation site.

amino acids 76-80, 92-96, 231-235, 289-293, 378-382, 421-425

Beta-transducin family Trp-Asp repeat protein.

amino acids 30-47, 105-118, 107-119, 203-216, 205-217, 296-308

FIGURE 265

TGGCCTCCCCAGCTTGCCAGGCACAAGGCTGAGCGGGAGGAAGCGAGAGGCATCTAAGCAGG
CAGTGTTTTGCCTTCACCCCAAGTGACCATGAGAGGTGCCACGCGAGTCTCAATCATGCTCC
TCCTAGTAACTGTGTCTGACTGTGCTGTGATCACAGGGGCCTGTGAGCGGGATGTCCAGTGT
GGGGCAGGCACCTGCTGTGCCATCAGCCTGTGGCTTCGAGGGCTGCGGATGTGCACCCCGCT
GGGGCGGGAAGGCGAGGAGTGCCACCCCGGCAGCCACAAGGTCCCCTTCTTCAGGAAACGCA
AGCACACACCTGTCCTTGCTTGCCCAACCTGCTGTGCTCCAGGTTCCCGGACGGCAGGTAC
CGCTGCTCCATGGACTTGAAGAACATCAATTTTTAGGCGCTTGCCTGGTCTCAGGATACCCA
CCATCCTTTTCCTGAGCACAGCCTGGATTTTTATTTCTGCCATGAAACCCAGCTCCCATGAC
TCTCCCAGTCCCTACACTGACTACCCTGATCTCTCTTGTCTAGTACGCACATATGCACACAG
GCAGACATACCTCCCATCATGACATGGTCCCCAGGCTGGCCTGAGGATGTCACAGCTTGAGG
CTGTGGTGTGAAAGGTGGCCAGCCTGGTTCTCTTCCCTGCTCAGGCTGCCAGAGAGGTGGTA
AATGGCAGAAAGGACATTCCCCCTCCCCTCCCCAGGTGACCTGCTCTCTTTCCTGGGCCCTG
CCCCTCTCCCCACATGTATCCCTCGGTCTGAATTAGACATTCCCTGGGCACAGGCTCTTGGGT
GCATTGCTCAGAGTCCCAGGTCTTGGCCTGACCCTCAGGCCCTTCACGTGAGGTCTGTGAGG
ACCAATTTGTGGGTAGTTCATCTTCCCTCGATTGGTTAACTCCTTAGTTTCAGACCACAGAC
TCAAGATTGGCTCTTCCCAGAGGGCAGCAGACAGTCACCCCAAGGCAGGTGTAGGGAGCCCA
GGGAGGCCAATCAGCCCCCTGAAGACTCTGGTCCCAGTCAGCCTGTGGCTTGTGGCCTGTGA
CCTGTGACCTTCTGCCAGAATTGTCATGCCTCTGAGGGCCCCCTCTTACCACACTTTACCAGT
TAACCACTGAAGCCCCCAATTTCCACAGCTTTTCCATTAAAATGCAAATGGTGGTGGTTCAA
TCTAATCTGATATTGACATATTAGAAGGCAATTAGGGTGTTTTCCTTAAACAACCTCCTTTCCA
AGGATCAGCCCTGAGAGCAGGTGGTGACTTTGAGGAGGGCAGTCCTCTGTCCAGATTGGGG
TGGGAGCAAGGGACAGGGAGCAGGGCAGGGCTGAAAGGGGCACTGATTGAGACCAGGGAGG
CAACTACACACCAACATGCTGGCTTTAGAATAAAAGCACCAACTGAAAAAA

266/330

FIGURE 266

MRGATRVSIMLLLVTVSDCAVITGACERDVQCGAGTCCAI SLWLRGLRMCTPLGREGE ECHP
GSHKVPFFRKRKHHTCPCLPNLLCSRFPDGRYRCSMDLKNINF

Signal peptide:

amino acids 1-19

Tyrosine kinase phosphorylation site:

amino acids 88-95

N-myristoylation sites:

amino acids 33-39, 35-41, 46-52

Downloaded from www.ashg.org

FIGURE 267

AGCGCCCGGGCGTCGGGGCGGTAAAAGGCCGGCAGAAGGGAGGCACTTGAGAAATGTCTTTT
CTCCAGGACCCAAGTTTCTTCACCATGGGGATGTGGTCCATTGGTGCAGGAGCCCTGGGGGC
TGCTGCCTTGGCATTGCTGCTTGCCAACACAGACGTGTTTCTGTCCAAGCCCCAGAAAGCGG
CCCTGGAGTACCTGGAGGATATAGACCTGAAAACACTGGAGAAGGAACCAAGGACTTTCAAA
GCAAAGGAGCTATGGGAAAAAATGGAGCTGTGATTATGGCCGTGCGGAGGCCAGGCTGTTT
CCTCTGTGAGAGGAAGCTGCGGATCTGTCTCCCTGAAAAGCATGTTGGACCAGCTGGGCG
TCCCCCTCTATGCAGTGGTAAAGGAGCACATCAGGACTGAAGTGAAGGATTTCCAGCCTTAT
TTCAAAGGAGAAATCTTCCTGGATGAAAAGAAAAAGTTCTATGGTCCACAAAGGCGGAAGAT
GATGTTTATGGGATTTATCCGTCTGGGAGTGTGGTACAACCTTCTTCCGAGCCTGGAACGGAG
GCTTCTCTGGAAACCTGGAAGGAGAAGGCTTCATCCTTGGGGGAGTTTTTCGTGGTGGGATCA
GGAAAGCAGGGCATTCTTCTTGAGCACCGAGAAAAAGAATTTGGAGACAAAGTAAACCTACT
TTCTGTTCTGGAAGCTGCTAAGATGATCAAACCACAGACTTTGGCCTCAGAGAAAAATGAT
TGTGTGAAACTGCCCAGCTCAGGGATAACCAGGGACATTCACCTGTGTTTATGGGATGTATT
GTTTCCACTCGTGTCCCTAAGGAGTGAGAAACCCATTTATACTCTACTCTCAGTATGGATTA
TTAATGTATTTTAATATTCTGTTTAGGCCCCACTAAGGCAAAAATAGCCCCAAAACAAGACTGA
CAAAAATCTGAAAAACTAATGAGGATTATTAAGCTAAAACCTGGGAAATAGGAGGCTTAAAA
TTGACTGCCAGGCTGGGTGCAGTGGCTCACACCTGTAATCCCAGCACTTTGGGAGGCCAAGG
TGAGCAAGTCACTTGAGGTGCGGAGTTCGAGACCAGCCTGAGCAACATGGCGAAACCCCGTC
TCTACTAAAAATACAAAATCACCCGGGTGTGGTGGCAGGCACCTGTAGTCCCAGCTACCCG
GGAGGCTGAGGCAGGAGAATCACTTGAACCTGGGAGGTGGAGGTTGCGGTGAGCTGAGATCA
CACCCTGTATTCCAGCCTGGGTGACTGAGACTCTAACTAA

FIGURE 268

MSFLQDPSFFTGMWWSIGAGALGAAALALLANTDVFLSKPQKALEYLEDIDLKTLEKEPR
TFKAKELWEKNGAVIMAVRRPGCFLCREEAADLSSLKSMLDQLGVPLYAVVKEHIRTEVKDF
QPYFKGEIFLDEKKKFYGPQRRKMMFMGFIRLGWYNFFRAWNGGFSGNLEGEFGFILGGVFV
VGSGKQGILLEHREKEFGDKVNLLSVLEAAKMIKPQTLASEKK

268/330

FIGURE 269

ACGGACCGAGGGTTCGAGGGAGGGACACGGACCAGGAACCTGAGCTAGGTCAAAGACGCCCCG
GGCCAGGTGCCCCGTCGCAGGTGCCCCTGGCCGGAGATGCGGTAGGAGGGGCGAGCGCGAGA
AGCCCCCTTCCTCGGGCGCTGCCAACCCGCCACCCAGCCCATGGCGAACCCCGGGCTGGGGCTG
CTTCTGGCGCTGGGCCTGCCGTTCTTGCTGGCCCGCTGGGGCCGAGCCTGGGGGCAAATACA
GACCACTTCTGCAAATGAGAATAGCACTGTTTTGCCTTCATCCACCAGCTCCAGCTCCGATG
GCAACCTGCGTCCGGAAGCCATCACTGCTATCATCGTGGTCTTCTCCCTCTTGGCTGCCTTG
CTCCTGGCTGTGGGGCTGGCACTGTTGGTGCGGAAGCTTCGGGAGAAGCGGCAGACGGAGGG
CACCTACCGGCCCAGTAGCGAGGAGCAGTTCTCCCATGCAGCCGAGGCCCGGGCCCCCTCAGG
ACTCCAAGGAGACGGTGCAGGGCTGCCTGCCCATCTAGGTCCCCTCTCCTGCATCTGTCTCC
CTTCATTGCTGTGTGACCTTGGGGAAAGGCAGTGCCCTCTCTGGGCAGTCAGATCCACCCAG
TGCTTAATAGCAGGGAAGAAGGTACTTCAAAGACTCTGCCCCTGAGGTCAAGAGAGGATGGG
GCTATTCACCTTTTATATATTTATATAAAATTAGTAGTGAGATGTAAAAAAAAAAAAAAAAAAAA

LIFE

270/330

FIGURE 270

MANPGLGLLLALGLPFLLARWGRAWGQIQTTSANENSTVLPSSTSSSSDGNLRPEAITAIIV
VFSLLAALLLAVGLALLVRKLREKRQTEGTYRPSSEEQFSHAAEARAPQDSKETVQGCLPI

270/330

FIGURE 271

AATATATCATCTATTTATCATTAATCAATAATGTATTCTTTTATTCCAATAACATTTGGGTT
TTGGGATTTTAATTTTCAAACACAGCAGAAATGACATTTTTTCTGTCACTATTATTATTGTTG
GTATGTGAAGCTATTTGGAGATCCAATTCAAGGAAGCAACACATTGGAGAATGGCTACTTTCT
ATCAAGAAATAAAGAGAACCACAGTCAACCCACACAATCATCTTTAGAAGACAGTGTGACTC
CTACCAAAGCTGTCAAACCACAGGCAAGGGCATAGTTAAAGGACGGAATCTTGACTCAAGA
GGGTAAATTCTTGGTGCTGAAGCCTGGGGCAGGGGTGTAAAGAAAAACACTTAGATTCAATG
ATTGTAAATTTAAGGCAAATACACATATTAGTATTACCTTAGTGTAATGTATCCCTGTCATA
TATACAATAAGGTGAAATTATAAGTACCCTATGCAGTTGGCTGGACAGTTCTAAATTGGACT
TTATTAATTTTAAAATCAGTAACTGATTTATCACTGGCTATGTGCTTAGATCTACAGGAGA
TCATATAATTTGATACAAATAAAAGAAAAGTGTCTCTCCCCTTACAGAATTGACATTTTAA
ATGCGATACAGTTAGAATAGGAAATATGACATTAGAAAGGAAGAATGACAGGGAGAAAGGAA
AGAAGGGAAAATGTTGCCAAGGAAAAAAAAA

271/330

272/330

FIGURE 272

MTFFLSLLLLLVCEAIWRSNSGSNTLENGYFLSRNKENHSQPTQSSLEDSVTPTKAVKTTGK
GIVKGRNLDSRGLILGAEAWGRGVKKNT

2025-11-14 14:14:14

FIGURE 273

GCCAGGAATAACTAGAGAGGAACA**ATG**GGGTTATTTCAGAGGTTTTGTTTTCTCTTAGTTCT
 GTGCCCTGCTGCACCAGTCAAATACTTCCTTCATTAAGCTGAATAATAATGGCTTTGAAGATA
 TTGTCATTGTTATAGATCCTAGTGTGCCAGAAGATGAAAAAATAATTGAACAAATAGAGGAT
 ATGGTGACTACAGCTTCTACGTACCTGTTTGAAGCCACAGAAAAAAGATTTTTTTTCAAAAA
 TGTATCTATATTAATTCTCTGAGAATTGGAAGGAAAATCCTCAGTACAAAAGGCCAAAACATG
 AAAACCATAAACATGCTGATGTTATAGTTGCACCACCTACACTCCCAGGTAGAGATGAACCA
 TACACCAAGCAGTTTACAGAATGTGGAGAGAAAGGCGAATACATTCACCTTCCCCCTGACCT
 TCTACTTGGAAAAAACAATAATGAATATGGACCACCAGGCAAACTGTTTGTCCATGAGTGGG
 CTCACCTCCGGTGGGGAGTGTGATGAGTACAATGAAGATCAGCCTTTCTACCGTGCTAAG
 TCAAAAAAATCGAAGCAACAAGGTGTTCCGCAGGTATCTCTGGTAGAAATAGAGTTTATAA
 GTGTCAAGGAGGCAGCTGTCTTAGTAGAGCATGCAGAATTGATTCTACAACAAAACCTGTATG
 GAAAAGATTGTCAATTCTTTCTGATAAAGTACAAACAGAAAAAGCATCCATAATGTTTATG
 CAAAGTATTGATTCTGTTGTTGAATTTTGTAAACGAAAAAACCCATAATCAAGAAGCTCCAAG
 CCTACAAAACATAAAGTGCAATTTTAGAAGTACATGGGAGGTGATTAGCAATCTGAGGATT
 TTA AAAACACCATAACCATGGTGACACCACCTCCTCCACCTGTCTTCTCATTGCTGAAGATC
 AGTCAAAGAATTGTGTGCTTAGTTCTTGATAAGTCTGGAAGCATGGGGGGTAAGGACCGCCT
 AAATCGAATGAATCAAGCAGCAAAACATTTCTGCTGCAGACTGTTGAAAATGGATCCTGGG
 TGGGGATGGTTCACCTTTGATAGTACTGCCACTATTGTAAATAAGCTAATCCAAATAAAAAGC
 AGTGATGAAAGAAACACACTCATGGCAGGATTACCTACATATCCTCTGGGAGGAACTTCCAT
 CTGCTCTGGAATTAAATATGCATTTTCAAGGTGATTGGAGAGCTACATTCCTCAACTCGATGGAT
 CCGAAGTACTGCTGCTGACTGATGGGGAGGATAACACTGCAAGTCTTGTATGATGAAGTG
 AAACAAAGTGGGGCCATTGTTCATTTTATTGCTTTGGGAAGAGCTGCTGATGAAGCAGTAAT
 AGAGATGAGCAAGATAACAGGAGGAAGTCATTTTTATGTTTCAGATGAAGCTCAGAACAATG
 GCCTCATTGATGCTTTTGGGGCTCTTACATCAGGAAATACTGATCTCTCCCAGAAGTCCCTT
 CAGCTCGAAAGTAAGGGATTAACACTGAATAGTAATGCCTGGATGAACGACACTGTCATAAT
 TGATAGTACAGTGGGAAAGGACACGTTCTTTCTCATCACATGGAACAGTCTGCCTCCCAGTA
 TTTCTCTCTGGGATCCAGTGGAAACAATAATGGAAAAATTTACAGTGGATGCAACTTCCAAA
 ATGGCCTATCTCAGTATTTCCAGGAAGTCAAAAGGTGGGCACTTGGGCATACAATCTTCAAGC
 CAAAGCGAACCCAGAAACATTAACATATTACAGTAACCTTCTCGAGCAGCAAAATCTTCTGTGC
 CTCCAATCACAGTGAATGCTAAAATGAATAAGGACGTAAACAGTTTCCCCAGCCCAATGATT
 GTTTACGCAGAAATTCTACAAGGATATGTACCTGTTCTTGGAGCCAATGTGACTGCTTTTAT
 TGAATCACAGAATGGACATACAGAAGTTTTTGAACCTTTTGGATAATGGTGCAGGCGCTGATT
 CTTTCAAGAATGATGGAGTCTACTCCAGGTATTTTACAGCATATACAGAAAATGGCAGATAT
 AGCTTAAAGTTTCGGGCTCATGGAGGAGCAAACTGCCAGGCTAAAATTACGGGCTCCACT
 GAATAGAGCCGCGTACATACCAGGCTGGGTAGTGAACGGGGAAATTAAGCAAACCCGCCAA
 GACCTGAAATTGATGAGGATACTCAGACCACCTTGGAGGATTTAGCCGAACAGCATCCGGA
 GGTGCATTTGTGGTATCACAGTCCCAAGCCTTCCCTTGCCTGACCAATACCCACCAAGTCA
 AATCACAGACCTTGATGCCACAGTTCATGAGGATAAGATTATTCTTACATGGACAGCACCAG
 GAGATAATTTTGTGTTGGAAAAGTTCACGTTATATCATAAGAATAAGTGCAAGTATTCTT
 GATCTAAGAGACAGTTTTGATGATGCTCTTCAAGTAAATACTACTGATCTGTCACCAAAGGA
 GGCCAACTCCAAGGAAAGCTTTGCATTTAAACCAGAAAAATATCTCAGAAGAAAATGCAACCC
 ACATATTTTATTGCCATTAAAAGTATAGATAAAAGCAATTTGACATCAAAAGTATCCAACATT
 GCACAAGTAACTTTGTTTATCCCTCAAGCAAATCCTGATGACATTGATCCTACACCTACTCC
 TACTCCTACTCCTACTCCTGATAAAAGTCATAATTCTGGAGTTAATATTTCTACGCTGGTAT
 TGTCTGTGATTGGGTCTGTTGTAATTGTTAACTTTATTTTAAAGTACCACCATT**TGA**ACCTTA
 ACGAAGAAAAAATCTTCAAGTAGACCTAGAAGAGAGTTTTTAAAAAACAAAAACAATGTAAGT
 AAAGGATATTTCTGAATCTTAAAATTCATCCCATGTGTGATCATAAACTCATAAAAAATAATT
 TTAAGATGTGGGAAAGGATACTTTGATTAAATAAAAAACACTCATGGATATGTAAAAACTGT
 CAAGATTAAAAATTAATAGTTTTCATTTATTTGTTATTTTATTTGTAAGAAATAGTGATGAAC
 AAAGATCCTTTTTCATACTGATACCTGGTTGTATATTATTTGATGCAACAGTTTTCTGAAAT
 GATATTTCAAATTGCAATCAAGAAATTAATAATCATCTATCTGAGTAGTCAAAATACAAGTAA
 GGAGAGCAAATAAACAAACATTTGGAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
 AA

FIGURE 274

MGLFRG FVFL LVLCLLHQSN TSFIKLN NNGFEDIVIVIDPSVPEDEKII EQIEDMVT TASTY
LFEATEKR FFFK NVSILIPENWK ENPQYKR PKH ENHKHADVIVAPPTLPGRDEPYTKQFTEC
GEKGEYIHF TPDLLLGKKQNEYGPPGKLFVHEWAHLRWGVFDEYNEDQPFYRAKSKKIEATR
CSAGISGRNRVYKCQGG SCLSRACRIDSTTKLYGKDCQFFPDKVQTEKASIMFMQSIDSVVE
FCNEKTHNQEAPSLQNIKCNFRSTWEVISNSEDFKNTIPMVT P P P P P V F S L L K I S Q R I V C L V
L D K S G S M G G K D R L N R M N Q A A K H F L L Q T V E N G S W V G M V H F D S T A T I V N K L I Q I K S S D E R N T L M
A G L P T Y P L G G T S I C S G I K Y A F Q V I G E L H S Q L D G S E V L L L T D G E D N T A S S C I D E V K Q S G A I V H
F I A L G R A D E A V I E M S K I T G G S H F Y V S D E A Q N N G L I D A F G A L T S G N T D L S Q K S L Q L E S K G L T
L N S N A W M N D T V I I D S T V G K D T F F L I T W N S L P P S I S L W D P S G T I M E N F T V D A T S K M A Y L S I P G
T A K V G T W A Y N L Q A K A N P E T L T I T V T S R A A N S S V P P I T V N A K M N K D V N S F P S P M I V Y A E I L Q G
Y V P V L G A N V T A F I E S Q N G H T E V L E L L D N G A G A D S F K N D G V Y S R Y F T A Y T E N G R Y S L K V R A H G
G A N T A R L K L R P P L N R A A Y I P G W V V N G E I E A N P P R P E I D E D T Q T T L E D F S R T A S G G A F V V S Q V
P S L P L P D Q Y P P S Q I T D L D A T V H E D K I I L T W T A P G D N F D V G K V Q R Y I I R I S A S I L D L R D S F D D
A L Q V N T T D L S P K E A N S K E S F A F K P E N I S E E N A T H I F I A I K S I D K S N L T S K V S N I A Q V T L F I P
Q A N P D D I D P T P T P T P T P D K S H N S G V N I S T L V L S V I G S V V I V N F I L S T T I

Signal peptide:

amino acids 1-21

Putative transmembrane domains:

amino acids 284-300, 617-633

Leucine zipper pattern.

amino acids 469-491, 476-498

N-glycosylation site.

amino acids 20-24, 75-79, 340-344, 504-508, 542-546, 588-592,
628-632, 811-815, 832-836, 837-841, 852-856, 896-900

FIGURE 275

CTCCTTAGGTTGAAACCCTGGGAGTAGAGTACTGACAGCAAAGACCGGGAAAGACCATACGTCCCCG
 GGCAGGGGTGACAACAGGTGTCTCTTTTGTCTCGTGTGTGGCTGCCTTCCTATTTCAAGGAAAG
 ACGCCAAGGTAATTTTGACCCAGAGGAGCAATGATGTAGCCACCTCCTAACCTTCCCTTCTTGAACC
 CCCAGTTATGCCAGGATTTACTAGAGAGTGTCAACTCAACCAGCAAGCGGCTCCTTCGGCTTAACCT
 GTGGTTGGAGGAGAGAACCTTTGTGGGGCTGCGTTCTCTTAGCAGTGCTCAGAAGTGACTTGCCTGA
 GGGTGGACCAGAAGAAAGGAAAGGTCCCCCTCTTGCTGTTGGCTGCACATCAGGAAGGCTGTGATGGG
 AATGAAGGTGAAACCTTGGAGATTTCACTTCAGTCTGCTTCTGCTGCAAGATCATCCTTTAAAA
 GTAGAGAAGTGTCTGTGTGGTGGTTAACTCCAAGAGGCAGAACTCGTTCTAGAAGGAAATGGATG
 CAAGCAGCTCCGGGGGCCCCAAACGCATGCTTCCTGTGGTCTAGCCCAGGGAAGCCCTTCCGTGGGG
 GCCCCGGCTTTGAGGGATGCCACCGGTTCTGGACGCATGGCTGATTCTGAAATGATGATGGTTCCGC
 GGGGGCTGCTTGCCTGGATTTCCCGGGTGGTGGTTTGTGCTGGTGTCTCCTCTGCTGTGCTATCTCTGT
 CCTGTACATGTTGGCCTGCACCCCAAAAGGTGACGAGGAGCAGCTGGCACTGCCAGGGCCAACAGC
 CCCACGGGGAAGGAGGGGTACCAGGCCGTCTTCAGGAGTGGGAGGAGCAGCACCGCAACTACGTGA
 GCAGCCTGAAGCGGCAGATCGCACAGCTCAAGGAGGAGCTGCAGGAGAGGAGTGAGCAGCTCAGGAA
 TGGGCAGTACCAAGCCAGCGATGCTGCTGGCTGGGTCTGGACAGGAGCCCCCAGAGAAAACCCAG
 GCCGACCTCCTGGCCTTCTGCACTCGCAGGTGGACAAGGCAGAGGTGAATGCTGGCGTCAAGCTGG
 CCACAGAGTATGCAGCAGTGCCTTTTCGATAGCTTTACTCTACAGAAGGTGTACCAGCTGGAGACTGG
 CCTTACCCGCCACCCGAGGAGAAGCCTGTGAGGAAGGACAAGCGGGATGAGTTGGTGGAAAGCCATT
 GAATCAGCCTTGGAGACCCTGAACAATCCTGCAGAGAACAGCCCCAATCACCGTCTTACACGGCCT
 CTGATTTTCATAGAAGGGATCTACCGAACAGAAAGGGACAAAGGGACATTGTATGAGCTCACCTTCAA
 AGGGGACCACAAAACGAATTCAAACGGCTCATCTTATTTTCGACCATTCAGCCCCATCATGAAAGTG
 AAAAAATGAAAAGCTCAACATGGCCAACACGCTTATCAATGTTATCGTGCCTCTAGCAAAAAGGGTGG
 ACAAGTTCGGCAGTTTCATGCAGAATTTTCAGGGAGATGTGCATTGAGCAGGATGGGAGAGTCCATCT
 CACTGTTGTTTACTTTTGGGAAAGAAGAAATAAATGAAGTCAAAGGAATACTTGAAAACACTTCCAAA
 GCTGCCAACTTCAGGAACCTTTACCTTCATCCAGCTGAATGGAGAATTTTCTCGGGGAAAGGCACTTG
 ATGTTGGAGCCCGCTTCTGGAAGGGAAGCAACGCTCCTTCTCTTTTCTGTGATGTGGACATCTACTT
 CACATCTGAATTCCTCAATACGTGTAGGCTGAATACACAGCCAGGGAAGAAGGTATTTTATCCAGTT
 CTTTTAGTCAGTACAATCCTGGCATAATATACGGCCACCATGATGCAGTCCCTCCCTTGGAAACAGC
 AGCTGGTCATAAAGAAGGAACTGGATTTTGGAGAGACTTTGGATTGAGGATTTGGGATGTCAGTATCG
 GTCAGACTTCATCAATATAGGTGGGTTTGTATCTGGACATCAAAGGCTGGGGCGGAGAGGATGTGCAC
 CTTTATCGCAAGTATCTCCACAGCAACCTCATAGTGGTACGGACGCCTGTGCGAGGACTCTTCCACC
 TCTGGCATGAGAAGCGCTGCATGGACGAGCTGACCCCGAGCAGTACAAGATGTGCATGCAGTCCAA
 GGCCATGAACGAGGCATCCACCGGCCAGCTGGGCTGCTGGTGTTCAGGCACGAGATAGAGGCTCAC
 CTTTCGCAACAGAAACAGAAGACAAGTAGCAAAAAAACATGAAGTCCAGAGAAGGATTGTGGGAGA
 CACTTTTTCTTTCTTTTGAATTAAGTGGCTGCAACAGAGAAAAGACTTCCATAAAGGAGC
 ACAAAGAATTTGGACTGATGGGTGAGAGATGAGAAGCCTCCGATTCTCTCTGTTGGGCTTTTAC
 AACAGAAATCAAAATCTCCGCTTTGCCTGCAAAAGTAACCCAGTTGCACCCCTGTGAAGTGTCTGACA
 AAGGCAGAATGCTTGTGAGATTATAAGCCTAATGGTGTGGAGGTTTTGATGGTGTGTACAATACT
 GAGACCTGTTGTTTGTGTGCTCATTGAAATATTCATGATTTAAGAGCAGTTTTGTAATAAATTCAT
 TAGCATGAAAGGCAAGCATATTTCTCCTCATATGAATGAGCCTATCAGCAGGGCTTAGTTTTCTAGG
 AATGCTAAAATATCAGAAGGCAGGAGAGGAGATAGGCTTATTATGATACTAGTGAGTACATTAAAGTA
 AAATAAATGGAGCCAGAAAAGAAAAGAAACCATAAATATCGTGTCTATTTTCCCAAGATTAACCA
 AAAATAATCTGCTTATCTTTTTTGGTTGTCCTTTTAACTGTCTCCGTTTTTTTCTTTTATTTAAAAAT
 GCACCTTTTTTTCCCTTGTGAGTTATAGTCTGCTTATTTAATTACCACTTTGCAAGCCTTACAAGAGA
 GCACAAGTTGGCCTACATTTTTTATATTTTTTAAAGAGATACTTTGAGATGCATTATGAGAAGTTTCA
 GTTCAAAGCATCAAATTGATGCCATATCCAAGGACATGCCAAATGCTGATTCTGTCAGGCACTGAAT
 GTCAGGCATTGAGACATAGGGAAGGAATGTTTGTACTAATACAGACGTACAGATACTTTCTCTGAA
 GAGTATTTTCGAAGAGGAGCAACTGAACACTGGAGGAAAAGAAAATGACACTTTCTGCTTTACAGAA
 AAGGAAACTCATTGAGACTGGTGATATCGTGATGACCTAAAAGTCAGAAAACCATTTTTCTCCTCA
 GAAGTAGGGACCGCTTCTTACCTGTTTAAATAAACCAAGTATACCGTGTGAACCAACATCTCT
 TTTCAAAACAGGGTGCTCCTCCTGGCTTGGCTTCCATAAGAAGAAATGGAGAAAAATATATATAT
 ATATATATATATTGTGAAAGATCAATCCATCTGCCAGAATCTAGTGGGATGGAAGTTTTTGTACAT
 GTTATCCACCCAGGCCAGGTGGAAGTAACTGAATTATTTTTTAAATTAAGCAGTTCTACTCAATCA
 CCAAGATGCTTCTGAAAATTGCATTTTATACCATTTCAAACTATTTTTTAAAAATAAATACAGTTA
 ACATAGAGTGGTTTCTTCATTTCATGTGAAAATTATTAGCCAGCACCAGATGCATGAGCTAATTATCT
 CTTTGAGTCCTTGCTTCTGTTTGTCTACAGTAAACTCATTTGTTTAAAGCTTCAAGAACATTCAAGC
 TGTTGGTGTGTTAAAAATGCATTGTATTGATTGTACTGGTAGTTTATGAAATTTAATTAAAAACAC
 AGGCCATGAATGGAAGGTGGTATTGCACAGCTAATAAATATGATTTGTGGATATGAA

FIGURE 276

MMVRRGLLAWISRVVLLVLLCCAISVLYMLACTPKGDEEQLALPRANSPTGKEGYQAVLQ
EWEEQHRNYVSSLKRQIAQLKEELQERSEQLRNGQYQASDAAGLGDRSPPEKTQADLLAFL
HSQVDKAEVNAGVKLATEYAAVPFDSFTLQKVYQLETGLTRHPEEKPV RKDKRDELVEAIES
ALETINNPAENSPNHRPYTASDFIEGIYRTERDKGTLYELTFKGDHKHEFKRLILFRPFSP
MKVKNEKLNMAN TLINVIVPLAKRVDKFRQFMQNFREMCIEQDGRVHLTVVYFGKEEINEVK
GILENTSKAANFRNFTFIQLNGEFSRGKGLDVGARFWKGSNVLLFFCDVDIYFTSEFLNTR
LNTQPGKKVFYPVLF SQYNPGIIYGHHDVPPLEQQLVIKKETGFWRDFGFGMTCQYRSDFI
NIGGFDLDIKGWGGEDVHLYRKYLHSNLIVVRTPVRGLFHLWHEKRCMDELTP EQYKMCMQS
KAMNEASHGQLGMLVFRHEIEAHLRKQKQKTSSKKT

276/330

FIGURE 277

GAAAGAATGTTGTGGCTGCTCTTTTTTCTGGTGACTGCCATTCATGCTGAACTCTGTCAACC
AGGTGCAGAAAATGCTTTTAAAGTGAGACTTAGTATCAGAACAGCTCTGGGAGATAAAGCAT
ATGCCTGGGATACCAATGAAGAATACCTCTTCAAAGCGATGGTAGCTTTCTCCATGAGAAAA
GTTCCCAACAGAGAAGCAACAGAAATTTCCCATGTCCTACTTTGCAATGTAACCCAGAGGGT
ATCATTCTGGTTTGTGGTTACAGACCCTTCAAAAAATCACACCCTTCCTGCTGTTGAGGTGC
AATCAGCCATAAGAATGAACAAGAACCGGATCAACAATGCCTTCTTTCTAAATGACCAAAC
CTGGAATTTTTTAAAAATCCCTTCACACTTGCACCACCCATGGACCCATCTGTGCCCATCTG
GATTATTATATTTGGTGTGATATTTTGCATCATCATAGTTGCAATTGCACTACTGATTTTAT
CAGGGATCTGGCAACGTAGAAGAAAGAACAAAGAACCATCTGAAGTGGATGACGCTGAAGAT
AAGTGTGAAAACATGATCACAATTGAAAATGGCATCCCCTCTGATCCCCCTGGACATGAAGGG
GGGCATATTAATGATGCCTTCATGACAGAGGATGAGAGGCTCACCCCTCTCTGAAGGGCTGT
TGTTCTGCTTCCTCAAGAAATTAAACATTTGTTTCTGTGTGACTGCTGAGCATCCTGAAATA
CCAAGAGCAGATCATATATTTTGTTTCACCATTCTTCTTTTGTAATAAATTTTGAATGTGCT
TGAAAGTGAAAAGCAATCAATTATACCCACCAACACCACTGAAATCATAAGCTATTCACGAC
TCAAAATATTCTAAAATATTTTCTGACAGTATAGTGTATAAATGTGGTCATGTGGTATTTG
TAGTTATTGATTTAAGCATTTTGTAGAAATAAGATCAGGCATATGTATATATTTTCACACTTC
AAAGACCTAAGGAAAAATAAATTTTCCAGTGGAGAATACATATAATATGGTGTAGAAATCAT
TGAAAATGGATCCTTTTTGACGATCACTTATATCACTCTGTATATGACTAAGTAAACAAAAG
TGAGAAGTAATTATTGTAAATGGATGGATAAAAATGGAATTACTCATATACAGGGTGAATT
TTATCCTGTTATCACACCAACAGTTGATTATATATTTTCTGAATATCAGCCCCTAATAGGAC
AATTCTATTTGTTGACCATTCTACAATTTGTAAAAGTCCAATCTGTGCTAACTTAATAAAG
TAATAATCATCTCTTTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAA

0993667.114494

278/330

FIGURE 278

MLWLLFFLVTAIHAE LCQPGAENAFKVRLSIRTALGDKAYAWDTNEEYLFKAMVAFSMRKVP
NREATEISHVLLCNVTQRVSFWFVVTDPSKNHTLPAVEVQSAIRMNKNRINNAFFLNDQTLE
FLKIPSTLAPPMDPSVPIWIIIFGVIFCIIIVAIALLILSGIWQRRRKNKEPSEVDDAEDKC
ENMITIENGIPSDPLDMKGGILMMPS

278/330

FIGURE 279

AACTCAAACCTCCTCTCTCTGGGAAAACGCGGTGCTTGCTCCTCCCGGAGTGGCCTTGGCAGG
GTGTTGGAGCCCTCGGTCTGCCCCGTCCGGTCTCTGGGGCCAAGGCTGGGTTTCCCTC**ATGT**
ATGGCAAGAGCTCTACTCGTGCGGTGCTTCTTCTCCTTGGCATAACAGCTCACAGCTCTTTGG
CCTATAGCAGCTGTGGAAATTTATACCTCCCGGTGCTGGAGGCTGTTAATGGGACAGATGC
TCGGTTAAAATGCACTTTCTCCAGCTTTGCCCCCTGTGGGTGATGCTCTAACAGTGACCTGGA
ATTTTCGTCCTCTAGACGGGGGACCTGAGCAGTTTGTATTCTACTACCACATAGATCCCTTC
CAACCCATGAGTGGGCGGTTTAAGGACCGGGTGTCTTGGGATGGGAATCCTGAGCGGTACGA
TGCCTCCATCCTTCTCTGGAAACTGCAGTTCGACGACAATGGGACATAACCTGCCAGGTGA
AGAACCCACCTGATGTTGATGGGGTGATAGGGGAGATCCGGCTCAGCGTCGTGCACACTGTA
CGCTTCTCTGAGATCCACTTCCTGGCTCTGGCCATTGGCTCTGCCTGTGCACTGATGATCAT
AATAGTAATTGTAGTGGTCCTCTTCCAGCATTACCGGAAAAAGCGATGGGCCGAAAGAGCTC
ATAAAGTGGTGGAGATAAAATCAAAGAAGAGGAAAGGCTCAACCAAGAGAAAAAGGTCTCT
GTTTATTTAGAAGACACAGAC**TAA**CAATTTTAGATGGAAGCTGAGATGATTTCCAAGAACAA
GAACCCTAGTATTTCTTGAAGTTAATGGAACTTTTCTTTGGCTTTTCCAGTTGTGACCCGT
TTTCCAACCAGTTCTGCAGCATATTAGATTCTAGACAAGCAACACCCCTCTGGAGCCAGCAC
AGTGCTCCTCCATATCACCAGTCATACACAGCCTCATTATTAAGGTCTTATTTAATTTCAGA
GTGTAAATTTTTTCAAGTGCTCATTAGGTTTTATAAACAAGAAGCTACATTTTGGCCCTTAA
GACACTACTTACAGTGTTATGACTTGTATACACATATATTGGTATCAAAGGGGATAAAAGCC
AATTTGTCTGTTACATTTCTTTTACGTATTTCTTTTAGCAGCACTTCTGCTACTAAAGTTA
ATGTGTTTACTCTCTTTCTTCCCACATTCTCAATTAAAAGGTGAGCTAAGCCTCCTCGGTG
TTTCTGATTAACAGTAAATCCTAAATTCAACTGTAAATGACATTTTATTTTTATGTCTC
TCCTTAACTATGAGACACATCTTGTTTTACTGAATTTCTTTCAATATTCAGGTGATAGATT
TTTGTCG

280/330

FIGURE 280

MYGKSSTRAVLLLLGIQLTALWPAAVEIYTSRVLEAVNGTDARLKCTFSSFAPVGDALTVT
WNFRPLDGGPEQFVFYYHIDPFQPMGRFKDRVSWDGNPERYDASILLWKLQFDDNGTYTCQ
VKNPPDVDGVIGEIRLSVVHTVRFSEIHFLALAIGSACALMIIIVIVVVLFFQHYRKKRWAER
AHKVVEIKSKEEERLNQEKKVSVYLEDTD

0953567.4444

FIGURE 281

GCATTTTGTCTGTGCTCCCTGATCTTCAGGTCACCACCATGAAGTTCTTAGCAGTCCTGGT
ACTCTTGGGAGTTTCCATCTTTCTGGTCTCTGCCCAGAATCCGACAACAGCTGCTCCAGCTG
ACACGTATCCAGCTACTGGTCCTGCTGATGATGAAGCCCCTGATGCTGAAACCACTGCTGCT
GCAACCACTGCGACCACTGCTGCTCCTACCACTGCAACCACCGCTGCTTCTACCACTGCTCG
TAAAGACATTCCAGTTTTACCCAAATGGGTGGGGATCTCCCGAATGGTAGAGTGTGTCCCTT
GAATGGAATCAGCTTGAGTCTTCTGCAATTGGTCACAACCTATTCATGCTTCCTGTGATTTC
ATCCAACCTACTTACCTTGCCTACGATATCCCCTTTATCTCTAATCAGTTTATTTTCTTTCAA
ATAAAAAATAACTATGAGCAACATAAAAAAAAAAAAAA

281/330

282/330

FIGURE 282

MKFLAVLVLLGVSI FLVSAQNPTTAAPADTYPATGPADDEAPDAETTAAATTATTAAPTAT
TAASTARKDIPVLPKWVGDL PNGRVCP

282/330

FIGURE 283

GGACTCTGAAGGTCCCAAGCAGCTGCTGAGGCCCCCAAGGAAGTGGTTCCAACCTTGGACCC
CTAGGGGTCTGGATTTGCTGGTTAACAAGATAACCTGAGGGCAGGACCCCATAGGGGA**ATGC**
TACCTCCTGCCCTTCCACCTGCCCTGGTGTTCACGGTGGCCTGGTCCCTCCTTGCCGAGAGA
GTGTCCTGGGTGAGGGACGCAGAGGACGCTCACAGACTCCAGCCCTTTGTTACCGAGAGGAC
ACTTGGAAGGTCCAGCGATGGTCCGGAGTCCACACACAGACTGGCGGCAGGGCAGGAGGGG
GACAGTTCTGTTGTGCTTGGTTGGACAGTAAGAGGGTCTTGGCCAGTCCAGGGTGGGGGGCG
GCAAACCTCCATAAAGAACCAGAGGGTCTGGGCCCCGGCCACAGAGTCATCTGCCCAGCTCCT
CTGCTGCTGGCCAGTGGGAGTGGCACGAGGTGGGGCTTTGTGCCAG**TAAA**ACCACAGGCTGG
ATTTGCCTGCGGGCCATGGTCCCTGTCTAGGGCAGCAATTCTCAACCTTCTTGCTCTCAGGA
CCCCAAAGAGCTTTCATTGTATCTATTGATTTTTTACCACATTAGCAATTAAAACCTGAGAAAT
GGGCCGGGCACGGTGGCTCACGCCTGTAATCCAGCACTTTGGGAGGCCGAGGCGGGTGGAT
CACCTGAGATCAGGAGTTCAAGACCAGCCTGGCCAACATGGTGAAACCTTGTCTACTAAAAA
TACAAAAAATTAGCCAGGCACAGTGGTGTGCACTGGTAGTCCCAGTTACTCGGGAGGCTGAG
GCAGGAAAATCGCTTGAACCCAGGAGGCGGACGTTGCGGTGAGCCGAGATCGCGCCGCTGAT
TCCAGCCTGGGCGACAAGAGTGAGACTCCATCTCACACA

284/330

FIGURE 284

MLPPALPPALVFTVAWSLLAERVSWVRDAEDAHRLQPFVTERTLGKVQRWSGVHTQTGGRAG
GGQFCCAWLDSKRVLASPGWGAANSIKNQRVWAPATESSAQLLCCWPVGVARGGALCQ

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100

FIGURE 285

GTC**ATG**CCAGTGCCTGCTCTGTGCCTGCTCTGGGCCCTGGCAATGGTGACCCGGCCTGCCTCA
GCGGCCCCCATGGGCGGCCAGAACTGGCACAGCATGAGGAGCTGACCCTGCTCTTCCATGG
GACCCTGCAGCTGGGCCAGGCCCTCAACGGTGTGTACAGGACCACGGAGGGACGGCTGACAA
AGGCCAGGAACAGCCTGGGTCTCTATGGCCGCACAATAGAACTCCTGGGGCAGGAGGTCAGC
CGGGGCCGGGATGCAGCCCAGGAACTTCGGGCAAGCCTGTTGGAGACTCAGATGGAGGAGGA
TATTCTGCAGCTGCAGGCAGAGGCCACAGCTGAGGTGCTGGGGGAGGTGGCCCAGGCACAGA
AGGTGCTACGGGACAGCGTGCAGCGGCTAGAAGTCCAGCTGAGGAGCGCCTGGCTGGGCCCT
GCCTACCGAGAATTTGAGGTCTTAAAGGCTCACGCTGACAAGCAGAGCCACATCCTATGGGC
CCTCACAGGCCACGTGCAGCGGCAGAGGCGGGAGATGGTGGCACAGCAGCATCGGCTGCGAC
AGATCCAGGAGAGACTCCACACAGCGGCGCTCCCAGCC**TGA**ATCTGCCTGGATGGAACCTGAG
GACCAATCATGCTGCAAGGAACACTTCCACGCCCCGTGAGGCCCTGTGCAGGGAGGAGCTG
CCTGTTCACTGGGATCAGCCAGGGCGCCGGGCCCCACTTCTGAGCACAGAGCAGAGACAGAC
GCAGGCGGGGACAAAGGCAGAGGATGTAGCCCCATTGGGGAGGGGTGGAGGAAGGACATGTA
CCCTTTCATGCCTACACACCCCTCATTAAAGCAGAGTCGTGGCATTTCAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAA

286/330

FIGURE 286

MPVPALCLLWALAMVTRPASAAPMGGPELAQHEELTLLFHGTLQLGQALNGVYRTTEGRLTK
ARNSLGLYGRTIELLGQEVSRGRDAAQELRASLLETQMEEDILQLQAEATAEVLGEVAQAQK
VLRDSVQRLEVQLRSAWLGPAYREFEVLKAHADKQSHILWALTGHVQRQRREMVAQQHRLRQ
IQERLHTAALPA

286/330

FIGURE 287

GGCAACATGGCTCAGCAGGCTTGCCCCAGAGCCATGGCAAAGAATGGACTTGTAATTTGCAT
 CCTGGTGATCACCTTACTCCTGGACCAGACCACCAGCCACACATCCAGATTAAAAGCCAGGA
 AGCACAGCAAACGTCGAGTGAGAGACAAGGATGGAGATCTGAAGACTCAAATTGAAAAGCTC
 TGGACAGAAGTCAATGCCTTGAAGGAAATTCAAGCCCTGCAGACAGTCTGTCTCCGAGGCAC
 TAAAGTTCACAAGAAATGCTACCTTGCTTCAGAAGGTTTGAAGCATTTCATGAGGCCAATG
 AAGACTGCATTTCCAAAGGAGGAATCCTGGTTATCCCCAGGAACTCCGACGAAATCAACGCC
 CTCCAAGACTATGGTAAAAGGAGCCTGCCAGGTGTCAATGACTTTTGGCTGGGCATCAATGA
 CATGGTCACGGAAGGCAAGTTTGTGACGTCAACGGAATCGCTATCTCCTTCCTCAACTGGG
 ACCGTGCACAGCCTAACGGTGGCAAGCGAGAAAACGTGTCTCTCTCCCAATCAGCTCAG
 GGCAAGTGGAGTGATGAGGCCTGTCGCAGCAGCAAGAGATACATATGCGAGTTCACCATCCC
 TAAATAGGTCTTTCTCCAATGTGTCTCCTCCAAGCAAGATTCATCATAACTTATAGGTTTCATGA
 TCTCTAAGATCAAGTAAAAATCATAATTTTACTTATTAAAAAATTGCAACACAAGATCAAT
 GTCCATAGCAATATGATAGCATCAGCCAATTTTGCTAACACATTTCTTTGGGATTTTGCCCT
 TCCTGGGGTATAGGGGATCAGAAATATTGATCCATGTGCACGCAGATAAAATGGCTTCTGCT
 AAACAGACTAAAATCTTTCTCTCTAGTCTTTCTCACTTGTACAAACCCAGTTTGTTTTCAA
 AAATCACAGTAGCAATGCAACTCATCACTCTAGAAAAGCAAGCTTAGGCTACCTGAAAGATT
 TTCCCTTGGAAGTTTAGCGTATGTTTGACTAACAAAAATTCCTACATCAGAGACTCTAGGT
 GCTATATAATCCAAAACTTTTCAGCCTGTTGCTCATTCTGTCCCATGCTGGCAATAATACC
 TTGTCAGCCCATTACCCTTATTTTGAATTGCTCCATCTCCTGGTGGGACTTGTATCTTGTCT
 GCCATATCAGAACACAAACCCCTGAAGAGGTTCTGATTTGATTTTTTTTTTTTCTTCATGCC
 TACCCTTTTTTTTGAAGTTTCCAGCCGCAATTTGAAATGAAATGACAAGGTGTATATTTGAT
 CAATTTTCATTCCCACCATTGCATTACAACCTCTAACTTAAATGGGTAACCCTAAGGCATAT
 CAAAGAAGCAGATTGCATGATAAACGGAAATAGAAAAAAGAACCTACATTTATTTTGCTTT
 AGCATCCTTACTCTCACCTTTTATGAGATTGAGAGTGGACTTACATTTTCTTTTACATTT
 TCGTATATTTATTTTTTTTAGCCATCATTATATGTTTAAGTCTATTATGGGCAACCAATCTT
 TGGAAGCTGAAAACCTGAATTTAAAGAATGCTATCTTGGAATTTGCATACGTCTGTGCAATT
 TTTTATTCTGCCTAGTGCTATTCTGCTTGTTTAACTAGATTGTACAAAATAACTTCATTGCT
 TAATATCAAATTACAAAGTTTAGACTTGGAGGGAAATGGGCTTTTGAAGCAAACAATTTT
 AAATATATTTTGTCTTCAAATAAATAGTGTTTAAACATTGAATGTGTTTTGTGAACAATAT
 CCCACTTTGCAAACCTTAACTACACATGCTTGGAATTAAGTTTTAGCTGTTTTCATTTGCTCA
 ATAATAAAGCCTGAATTCTGATCAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

288/330

FIGURE 288

MAQQACPRAMAKNGLVICILVITLLLDQTTSHTSRLKARKHSKRRVRDKDGDLDKTQIEKLWT
EVNALKEIQALQTVCLRGTKVHKKCYLASEGLKHFHEANEDCISKGGILVIPRNSDEINALQ
DYGKRSLPGVNDFWLGINDMVTEGKFVDVNGIAISFLNWDRAQPNGGKRENCVLFSQSAQGK
WSDEACRSSKRYICEFTIPK

bioRxiv preprint doi: <https://doi.org/10.1101/2020.03.10.331111>; this version posted March 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 289

GCGAGGACCGGGTATAAGAAGCCTCGTGGCCTTGCCCGGGCAGCCGCAGGTTCCCCGCGCGC
CCCGAGCCCCCGCGCC**ATGA**AAGCTCGCCGCCCTCCTGGGGCTCTGCGTGGCCCTGTCCTGCA
GCTCCGCTGCTGCTTTCTTAGTGGGCTCGGCCAAGCCTGTGGCCCAGCCTGTCGCTGCGCTG
GAGTCGGCGGGCGGAGGCCGGGGCCGGGACCCTGGCCAACCCCTCGGCACCCTCAACCCGCT
GAAGCTCCTGCTGAGCAGCCTGGGCATCCCCGTGAACCACCTCATAGAGGGCTCCCAGAAGT
GTGTGGCTGAGCTGGGTCCCCAGGCCGTGGGGGCCGTGAAGGCCCTGAAGGCCCTGCTGGGG
GCCCTGACAGTGTTTGGCT**TGA**GCCGAGACTGGAGCATCTACACCTGAGGACAAGACGCTGCC
CACCCGCGAGGGCTGAAAACCCCGCCGCGGGGAGGACCGTCCATCCCCTTCCCCGGCCCCCT
CTCAATAAACGTGGTTAAGAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAA

Figure 289

290/330

FIGURE 290

MKLAALLGLCVALLSCSSAAFLVGSAPVAPVAALESAAEAGAGTLANPLGTNLNPLKLLLS
SLGIPVNHLEGSQKCVNELGPQAVGAVKALKALLGALTTFG

290/330

FIGURE 292

MKVVP S L L L S V L L A Q V W L V P G L A P S P Q S P E T P A P Q N Q T S R V V Q A P R E E E E D E Q E A S E E K A G E
E E K A W L M A S R Q Q L A K E T S N F G F S L L R K I S M R H D G N M V F S P F G M S L A M T G L M L G A T G P T E T Q I
K R G L H L Q A L K P T K P G L L P S L F K G L R E T L S R N L E L G L S Q G S F A F I H K D F D V K E T F F N L S K R Y F
D T E C V P M N F R N A S Q A K R L M N H Y I N K E T R G K I P K L F D E I N P E T K L I L V D Y I L F K G K W L T P F D P
V F T E V D T F H L D K Y K T I K V P M M Y G A G K F A S T F D K N F R C H V L K L P Y Q G N A T M L V V L M E K M G D H L
A L E D Y L T T D L V E T W L R N M K T R N M E V F F P K F K L D Q K Y E M H E L L R Q M G I R R I F S P F A D L S E L S A
T G R N L Q V S R V L R R T V I E V D E R G T E A V A G I L S E I T A Y S M P P V I K V D R P F H F M I Y E E T S G M L L F
L G R V V N P T L L

292/330

FIGURE 293

CTGGGATCAGCCACTGCAGCTCCCTGAGCACTCTCTACAGAGACGCGGACCCCAGACATGAG
GAGGCTCCTCCTGGTCACCAGCCTGGTGGTTGTGCTGCTGTGGGAGGCAGGTGCAGTCCCAG
CACCCAAGGTCCCTATCAAGATGCAAGTCAAACACTGGCCCTCAGAGCAGGACCCAGAGAAG
GCCTGGGGCGCCCGTGTGGTGGAGCCTCCGGAGAAGGACGACCAGCTGGTGGTGTCTGTTCCC
TGTCCAGAAGCCGAAACTCTTGACCACCGAGGAGAAGCCACGAGGTGAGGGCAGGGGCCCCA
TCCTTCCAGGCACCAAGGCCTGGATGGAGACCGAGGACACCCTGGGCCGTGTCCTGAGTCCC
GAGCCCGACCATGACAGCCTGTACCACCCCTCCGCCTGAGGAGGACCAGGGCGAGGAGAGGCC
CCGGTTGTGGGTGATGCCAAATCACCAGGTGCTCCTGGGACCGGAGGAAGACCAAGACCACA
TCTACCACCCCCAGTAGGGCTCCAGGGGCCATCACTGCCCCCGCCCTGTCCCAAGGGCCAGG
CTGTTGGGACTGGGACCCTCCCTACCCTGCCCCAGCTAGACAAATAAACCCCAGCAGGCAAA
AAAAAAAAAAAAAAAAA

GAGGCTCCTCCTGGTCACCAGCCTGGTGGTTGTGCTGCTGTGGGAGGCAGGTGCAGTCCCAG

FIGURE 295

AGAAAGCTGCACTCTGTTGAGCTCCAGGGCGCAGTGGAGGGAGGGAGTGAAGGAGCTCTCTG
TACCCAAGGAAAGTGCAGCTGAGACTCAGACAAGATTACA**ATG**AACCAACTCAGCTTCCTGC
TGTTTCTCATAGCGACCACCAGAGGATGGAGTACAGATGAGGCTAATACTTACTTCAAGGAA
TGGACCTGTTCTTCGTCTCCATCTCTGCCCAGAAGCTGCAAGGAAATCAAAGACGAATGTCC
TAGTGCATTTGATGGCCTGTATTTTCTCCGCACTGAGAATGGTGTATCTACCAGACCTTCT
GTGACATGACCTCTGGGGGTGGCGGCTGGACCCTGGTGGCCAGCGTGCATGAGAATGACATG
CGTGGGAAGTGCACGGTGGGCGATCGCTGGTCCAGTCAGCAGGGCAGCAAAGCAGACTACCC
AGAGGGGGACGGCAACTGGGCCAACTACAACACCTTTGGATCTGCAGAGGCGGCCACGAGCG
ATGACTACAAGAACCCTGGCTACTACGACATCCAGGCCAAGGACCTGGGCATCTGGCACGTG
CCCAATAAGTCCCCCATGCAGCACTGGAGAAACAGCTCCCTGCTGAGGTACCGCACGGACAC
TGGCTTCCTCCAGACACTGGGACATAATCTGTTTGGCATCTACCAGAAATATCCAGTGAAAT
ATGGAGAAGGAAAGTGTTGGACTGACAACGGCCCCGGTGATCCCTGTGGTCTATGATTTTGGC
GACGCCCAGAAAACAGCATCTTATTACTCACCTATGGCCAGCGGGAATTCAGTGCGGGATT
TGTTCAAGTTCAGGGTATTTAATAACGAGAGAGCAGCCAACGCCTTGTGTGCTGGAATGAGGG
TCACCGGATGTAACACTGAGCATCACTGCATTGGTGGAGGAGGATACTTTCCAGAGGCCAGT
CCCCAGCAGTGTGGAGATTTTCTGGTTTTGATTGGAGTGGATATGGAACATGTTGGTTA
CAGCAGCAGCCGTGAGATAACTGAGGCAGCTGTGCTTCTATTCTATCGT**TGA**GAGTTTTGTG
GGAGGGAACCCAGACCTCTCCTCCCAACCATGAGATCCCAAGGATGGAGAACAACCTTACCCA
GTAGCTAGAATGTTAATGGCAGAAGAGAAAACAATAAATCATATTGACTCAAGAAAAAAA

FIGURE 296

MNQLSFLLFLIATTRGWSTDEANTYFKEWTCSSSPSLPRSCKEIKDECPSAFDGLYFLRTEN
GVIYQTFCDMTSGGGGWTLVASVHENDMRGKCTVGDRWSSQQGSKADYPEGDGNWANYNTFG
SAEAATSDDYKNPGYYDIQAKDLGIWHVPNKSPMQHWRNSSLLRYRTDTGFLQTLGHNLFGL
YQKYPVKYGEKGCWTDNGPVI PVVYDFGDAQKTASYYSYPYGQREFTAGFVQFRVFNNERAAN
ALCAGMRVTGCNTEHHCIGGGGYFPEASPQQCGDFSGFDWSGYGTHVGYSRSSREITEAAVLL
FYR

296/330

FIGURE 297

GCGGAGCCGGCGCCGGCTGCGCAGAGGAGCCGCTCTCGCCGCGCCACCTCGGCTGGGAGCC
CACGAGGCTGCCGCATCCTGCCCTCGGAACAATGGGACTCGGCGCGCGAGGTGCTTGGGCCG
CGCTGCTCCTGGGGACGCTGCAGGTGCTAGCGCTGCTGGGGGCGCCCATGAAAGCGCAGCC
ATGGCGGCATCTGCAAACATAGAGAATTCTGGGCTTCCACACAACCTCCAGTGCTAACTCAAC
AGAGACTCTCCAACATGTGCCTTCTGACCATACAAATGAACTTCCAACAGTACTGTGAAAC
CACCAACTTCAGTTGCCTCAGACTCCAGTAATACAACGGTCACCACCATGAAACCTACAGCG
GCATCTAATACAACAACACCAGGGATGGTCTCAACAAATATGACTTCTACCACCTTAAAGTC
TACACCCAAAACAACAAGTGTTTCACAGAACACATCTCAGATATCAACATCCACAATGACCG
TAACCCACAATAGTTCAGTGACATCTGCTGCTTCATCAGTAACAATCACAACAACCTATGCAT
TCTGAAGCAAAGAAAGGATCAAAATTTGATACTGGGAGCTTTGTTGGTGGTATTGTATTAAC
GCTGGGAGTTTTATCTATTCTTTACATTGGATGCAAATGTATTACTCAAGAAGAGGCATTC
GGTATCGAACCATAGATGAACATGATGCCATCATTTAAGGAAATCCATGGACCAAGGATGGA
ATACAGATTGATGCTGCCCTATCAATTAATTTTGGTTTATTAATAGTTTAAACAATATTCT
CTTTTTGAAAATAGTATAAACAGGCCATGCATATAATGTACAGTGTATTACGTAAATATGTA
AAGATTCTTCAAGGTAACAAGGGTTTGGGTTTTGAAATAAACATCTGGATCTTATAGACCGT
TCATACAATGGTTTTAGCAAGTTCATAGTAAGACAAACAAGTCCTATCTTTTTTTTTTGGCT
GGGGTGGGGGCATTGGTCACATATGACCAGTAATTGAAAGACGTCATCACTGAAAGACAGAA
TGCCATCTGGGCATACAAATAAGAAGTTTGTCACAGCACTCAGGATTTTGGGTATCTTTTGT
AGCTCACATAAAGAACTTCAGTGCTTTTCAGAGCTGGATATATCTTAATTACTAATGCCACA
CAGAAATTATACAATCAAACCTAGATCTGAAGCATAATTTAAGAAAAACATCAACATTTTTTG
TGCTTTAACTGTAGTAGTTGGTCTAGAAACAAAATACTCC

298/330

FIGURE 298

MGLGARGAWAALLLGTLQVLALLGAAHESAAMAASANIENSGLPHNSSANSTETLQHVPSDH
TNETSNSTVKPPTSVASDSSNTTVTTMKPTAASNTTTPGMVSTNMTSTTLKSTPKTTSVSQN
TSQISTSTMTVTHNSSVTSAASSVTITTTMHSEAKKGSKFDTGSFVGGIVLT LGVLSILYIG
CKMYYSRRGIRYRTIDEHDAII

298/330

FIGURE 299

CAGCCGGGTCCCAAGCCTGTGCCTGAGCCTGAGCCTGAGCCTGAGCCCGAGCCGGGAGCCGG
TCGCGGGGGCTCCGGGCTGTGGGACCGCTGGGCCCCCAGCGATGGCGACCCCTGTGGGGAGGC
CTTCTTCGGCTTGGCTCCTTGCTCAGCCTGTGCTGCCTGGCGCTTTCCGTGCTGCTGCTGGC
GCAGCTGTCAGACGCCGCCAAGAATTTTCGAGGATGTCAGATGTAAATGTATCTGCCCTCCCT
ATAAAGAAAATTCTGGGCATATTTATAATAAGAACATATCTCAGAAAGATTGTGATTGCCTT
CATGTTGTGGAGCCCATGCCTGTGCGGGGGCCTGATGTAGAAGCATACTGTCTACGCTGTGA
ATGCAAATATGAAGAAAGAAGCTCTGTCACAATCAAGGTTACCATTATAATTTATCTCTCCA
TTTTGGGCCTTCTACTTCTGTACATGGTATATCTTACTCTGGTTGAGCCCATACTGAAGAGG
CGCCTCTTTGGACATGCACAGTTGATACAGAGTGATGATGATATTGGGGATCACCAGCCTTT
TGCAAATGCACACGATGTGCTAGCCCGCTCCCGCAGTCGAGCCAACGTGCTGAACAAGGTAG
AATATGCACAGCAGCGCTGGAAGCTTCAAGTCCAAGAGCAGCGAAAGTCTGTCTTTGACCGG
CATGTTGTCCTCAGCTAAATTGGGAATTGAATTCAAGGTGACTAGAAAGAAACAGGCAGACAA
CTGGAAAGAACTGACTGGGTTTTGCTGGGTTTTCATTTTAATACCTTGTTGATTTCACCAACT
GTTGCTGGAAGATTCAAACTGGAAGCAAAACTTGCTTGATTTTTTTTTCTTGTTAACGTA
ATAATAGAGACATTTTTAAAAGCACACAGCTCAAAGTCAGCCAATAAGTCTTTTCCCTATTTG
TGACTTTTACTAATAAAAATAAATCTGCCTGTAAATTATCTTGAAGTCCTTTACCTGGAACA
AGCACTCTCTTTTTTACCACATAGTTTTAACTTGACTTTCAAGATAATTTTCAGGGTTTTTG
TTGTTGTTGTTTTTTGTTTGTGTTTGGTGGGAGAGGGGAGGGATGCCTGGGAAGTGTT
AACAACTTTTTTCAAGTCACTTTACTAAACAACTTTTGTAATAGACCTTACCTTCTATTT
TCGAGTTTTATTTATATTTTGCAGTGTAGCCAGCCTCATCAAAGAGCTGACTTACTCATTTG
ACTTTTGCACTGACTGTATTATCTGGGTATCTGCTGTGTCTGCACTTCATGGTAAACGGGAT
CTAAATGCCTGGTGGCTTTTCACAAAAGCAGATTTTCTTCATGTACTGTGATGTCTGATG
CAATGCATCCTAGAACAACTGGCCATTTGCTAGTTTACTCTAAAGACTAAACATAGTCTTG
GTGTGTGTGGTCTTACTCATCTTCTAGTACCTTTAAGGACAAATCCTAAGGACTTGGA
TGCAATAAAGAAATTTTATTTTAAACCCAAGCCTCCCTGGATTGATAATATATACACATTTG
TCAGCATTTCCGGTCGTGGTGAGAGGCAGCTGTTTGAAGCTCCAATATGTGCAGCTTTGAACT
AGGGCTGGGGTTGTGGGTGCCTCTTCTGAAAGGTCTAACCATTATTGGATAACTGGCTTTTT
TCTTCCATATGTCCTCTTTGGAATGTAACAATAAAAATAATTTTTGAAACATCAA

300/330

FIGURE 300

MATLWGGLRLGSLLSLSCLALSVLLLAQLSDAAKNFEDVRCKCICPPYKENS
GHIYNKNIS
QKDCDCLHVVEPMPVRGPDVEAYCLRCECKYEERSSVTIKVTIIIIYLSILGLLLLYM
VYLTL
VEPILKRRLFGHAQLIQSDDDIGDHQPFANAHDVLA
RSRANVLNKVEYAQQRWKLQVQEQ
RKSVFDRHVVLS

300/330

301/330

FIGURE 301

GCACCTGCGACCACCGTGAGCAGTCATGGCGTACTCCACAGTGCAGAGAGTCGCTCTGGCTT
CTGGGCTTGTCTGGCTCTGTCGCTGCTGCTGCCCCAAGGCCTTCCTGTCCCGCGGGAAGCGG
CAGGAGCCGCGCCGACACCTGAAGGAAAATTGGGCCGATTTCACCTATGATGCATCATCA
CCAGGCACCCTCAGATGGCCAGACTCCTGGGGCTCGTTTCCAGAGGTCTCACCTTGCCGAGG
CATTTGCAAAGGCCAAAGGATCAGGTGGAGGTGCTGGAGGAGGAGGTAGTGGAAGAGGTCTG
ATGGGGCAGATTATTCCAATCTACGGTTTTGGGATTTTTTTATATATACTGTACATTCTATT
TAAGGTAAGTAGAATCATCCTAATCATATTACATCAATTGAAAATCTAATATGGCGATAAAAA
TCATTGTCTACATTAAACTTCTTATAGTTCATAAAATTATTTCAAATCCATCATCTCTTTA
AATCCTGCCTCCTCTTCATGAGGTACTTAGGATAGCCATTATTTTCAGTTTCACATAAGAATG
TTTACTCAATGTTTAAGTGTTTTGCCCCAAAATTCACAATAACAAGGCAGAACTAGGACTT
GAACATGGATCTTTTGTTCTTAATCCAGTGAGTGATACAATTCAATGCACTCCCCTGCCA

301/330

FIGURE 303

CGGCTCGAGTGCAGCTGTGGGGAGATTTTCAGTGCATTGCCTCCCCTGGGTGCTCTTCATCTT
 GGATTTGAAAGTTGAGAGCAGCATGTTTTGCCCACTGAAACTCATCCTGCTGCCAGTGTTAC
 TGGATTATTCCTTGGGCCTGAATGACTTGAATGTTTCCCCGCCTGAGCTAACAGTCCATGTG
 GGTGATTCACTCTGATGGGATGTGTTTTCCAGAGCACAGAAGACAAATGTATATTCAAGAT
 AGACTGGACTCTGTCACCAGGAGAGCACGCCAAGGACGAATATGTGCTATACTATTACTCCA
 ATCTCAGTGTGCCTATTGGGCGCTTCCAGAACCGCGTACACTTGATGGGGGACATCTTATGC
 AATGATGGCTCTCTCCTGCTCCAAGATGTGCAAGAGGCTGACCAGGGAACTATATCTGTGA
 AATCCGCCTCAAAGGGGAGAGCCAGGTGTTCAAGAAGGCGGTGGTACTGCATGTGCTTCCAG
 AGGAGCCCCAAAGAGCTCATGGTCCATGTGGGTGGATTGATTCAAGATGGGATGTGTTTTCCAG
 AGCACAGAAGTGAACACGTGACCAAGGTAGAATGGATATTTTCAGGACGGCGCGCAAAGGA
 GGAGATTGTATTTTCGTTACTACCACAACTCAGGATGTCTGTGGAGTACTCCCAGAGCTGGG
 GCCACTTCCAGAATCGTGTGAACCTGGTGGGGGACATTTTCCGCAATGACGGTTCATCATG
 CTTCAAGGAGTGAGGGAGTCAGATGGAGGAACTACACCTGCAGTATCCACCTAGGGAACCT
 GGTGTTCAAGAAAACCATTTGTGCTGCATGTCAGCCCGGAAGAGCCTCGAACACTGGTGACCC
 CGGCAGCCCTGAGGCCTCTGGTCTTGGGTGGTAATCAGTTGGTGATCATTGTGGGAATTGTC
 TGTGCCACAATCCTGCTGCTCCCTGTTCTGATATTGATCGTGAAGAAGACCTGTGGAAATAA
 GAGTTCAGTGAATTCTACAGTCTTGGTGAAGAACACGAAGAAGACTAATCCAGAGATAAAAG
 AAAAACCTGCCATTTTGAAAGATGTGAAGGGGAGAAACACATTTACTCCCCAATAATTGTA
 CGGGAGGTGATCGAGGAAGAAGAACCAAGTGAAAAATCAGAGGCCACCTACATGACCATGCA
 CCCAGTTTGGCCTTCTCTGAGGTGAGATCGGAACAACCTCACTTGAAAAAAGTCAGGTGGGG
 GAATGCCAAAAACACAGCAAGCCTTTTTGAGAGAAGATGGAGAGTCCCTTCATCTCAGCAGCGG
 TGGAGACTCTCTCCTGTGTGTGTCCTGGGCCACTCTACCAGTGATTTTCAGACTCCCGCTCTC
 CCAGCTGTCTCCTGTCTCATTGTTTGGTCAATACACTGAAGATGGAGAATTTGGAGCCTGG
 CAGAGAGACTGGACAGCTCTGGAGGAACAGGCCTGCTGAGGGGAGGGGAGCATGGACTTGGC
 CTCTGGAGTGGGACACTGGCCCTGGGAACCAGGCTGAGCTGAGTGGCCTCAAACCCCCCGTT
 GGATCAGACCCTCCTGTGGGCAGGGTTCTTAGTGGATGAGTTACTGGGAAGAATCAGAGATA
 AAAACCAACCCAAATCAA

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

FIGURE 304

MFCPLKLILLPVLLDYSLGLNDLNVSPPELTVHVGDSALMGCVFQSTEDKCIFKIDWTLSPG
EHAKDEYVLYYYSNLSVPIGRFQNRVHLMGDILCNDGSLLLQDVQEADQGTYYICEIRLKGES
QVFKKAVVLHVLPEEPKELMVHVGGLIQMGCVFQSTEVKHVTKVEWIFSGRRAKEEIVFRYY
HKLRMSVEYSQSWGHEFQNRVNLVGDI FRNDGSIMLQGVRES DGGNYTCSIHLGNLVFKKTIV
LHVSPEEPRTLVT PAALRPLVLGGNQLV IIVGIVCATILLLPVLILIVKKTCGNKSSVNSTV
LVKNTKKTNPEIKEKPCHEFERCEGEKHIYSPIIVREVIEEEEPSEKSEATYMTMHPVWPSLR
SDRNNNSLEKKSGGGMPKTQQAF

304/330

FIGURE 307

CCCACGCGTCCGCGCAGTCGCGCAGTTCTGCCTCCGCCTGCCAGTCTCGCCCGCGATCCCGG
CCCGGGGCTGTGGCGTCGACTCCGACCCAGGCAGCCAGCAGCCCGCGCGGGAGCCGGACCGC
CGCCGGAGGAGCTCGGACGGCATGCTGAGCCCCCTCCTTTGCTGAAGCCCGAGTGCGGAGAA
GCCCCGGGCAAACGCAGGCTAAGGAGACCAAAGCGGCGAAGTCGCGAGACAGCGGACAAGCAG
CGGAGGAGAAGGAGGAGGAGGCGAACCCAGAGAGGGGCGAGCAAAAGAAGCGGTGGTGGTGGG
CGTCGTGGCC**ATG**GCGGCGGCTATCGCCAGCTCGCTCATCCGTCAGAAGAGGCAAGCCCGCG
AGCGCGAGAAATCCAACGCCTGCAAGTGTGTCAGCAGCCCCAGCAAAGGCAAGACCAGCTGC
GACAAAAACAAGTTAAATGTCTTTCCCGGGTCAAACCTCTTCGGCTCCAAGAAGAGGGCGCAG
AAGAAGACCAGAGCCTCAGCTTAAGGGTATAGTTACCAAGCTATACAGCCGACAAGGCTACC
ACTTGCAGCTGCAGGCGGATGGAACCATTTGATGGCACCAAAGATGAGGACAGCACTTACACT
CTGTTTAACTCATCCCTGTGGGTCTGCGAGTGGTGGCTATCCAAGGAGTTCAAACCAAGCT
GTACTTGGCAATGAACAGTGAGGGATACTTGTACACCTCGGAACCTTTTCACACCTGAGTGCA
AATTCAAAGAATCAGTGTTTGAAAATTATTATGTGACATATTCATCAATGATATACCGTCAG
CAGCAGTCAGGCCGAGGGTGGTATCTGGGTCTGAACAAAGAAGGAGAGATCATGAAAGGCAA
CCATGTGAAGAAGAACAAGCCTGCAGCTCATTTTCTGCCTAAACCACTGAAAGTGGCCATGT
ACAAGGAGCCATCACTGCACGATCTCACGGAGTTCTCCCGATCTGGAAGCGGGACCCCAACC
AAGAGCAGAAGTGTCTCTGGCGTGCTGAACGGAGGCAAATCCATGAGCCACAATGAATCAAC
GTAGCCAGTGAGGGCAAAGAAGGGCTCTGTAAACAGAACCTTACCTCCAGGTGCTGTTGAAT
TCTTCTAGCAGTCCTTCACCCAAAAGTTCAAATTTGTCAGTGACATTTACCAAACAAACAGG
CAGAGTTCACTATTCTATCTGCCATTAGACCTTCTTATCATCCATACTAAAGC

FIGURE 308

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA28498

><subunit 1 of 1, 245 aa, 1 stop

><MW: 27564, pI: 10.18, NX(S/T): 1

MAAAIASSLIRQKRQAREREKSNACKCVSSPSKGKTSCDKNKLVFSRVKLFSGSKRRRRRP
EPQLKGIVTKLYSRQGYHLQLQADGTIDGTKDEDSTYTLFNLIPVGLRVVAIQGVQTKLYLA
MNSEGILYTSELFTECKFKESVFENYYVTYSSMIYRQQQSGRGWYLGLENKEGEIMKGNHVK
KNKPAAHFLPKPLKVAMYKEPSLHDLTEFSRSGSGTPTKSRSVSGVLNGGKSMHNEST

N-glycosylation site.

amino acids 242-246

Glycosaminoglycan attachment site.

amino acids 165-169, 218-222

Tyrosine kinase phosphorylation site.

amino acids 93-100

N-myristoylation site.

amino acids 87-93, 231-237

ATP/GTP-binding site motif A (P-loop).

amino acids 231-239

HBGF/FGF family proteins

amino acids 78-94, 102-153

FIGURE 310

MAALASSLIRQKREVREPGGSRPVSAQRRVCPRGTKSLCQKQLLILLSKVRLCGGRPARPDR
GPEPQLKGIVTKLFCRQGFYLAQNPDGSIQGTPEDTSSFTHFNLIPVGLRVVTIQSAKLGHY
MAMNAEGLLYSSPHFTAECRFKECVFENYYVLYASALYRQRRSGRAWYLGLDKEGQVMKGNR
VKKTKAAAHFLPKLLEVAMYQEPSLHSVPEASPSSPPAP

Tyrosine kinase phosphorylation site:

amino acids 199-207

N-myristoylation sites:

amino acids 54-60, 89-95, 131-137

HBGF/FGF family signature:

amino acids 131-155

310/330

ATAG

FIGURE 312

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA28503

><subunit 1 of 1, 247 aa, 1 stop

><MW: 27702, pI: 10.36, NX(S/T): 2

MAAAIASGLIRQKRQAREQHWD RPSASRRRSSPSKNRGLCNGNLVDIFSKVRI FGLKKRRLR
RQDPQLKGIVTRLYCRQGYYLQMH PDGALDGTKDDSTNSTLFNLIPVGLRVVAIQGVKTGLY
IAMNGEGYLYPSELF TPECKFKESVFENYYVIYSSMLYRQQESGRAWFLGLNKEGQAMKGNR
VKKTKPAAHFLPKPLEVAMYREPSLHDVGETVPKPGVTPSKSTSASAIMNGGKPVNKS KTT

N-glycosylation site.

amino acids 100-104, 242-246

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 28-32, 29-33

Tyrosine kinase phosphorylation site.

amino acids 199-207

N-myristoylation site.

amino acids 38-44, 89-95, 118-124, 122-128, 222-228

HBGF/FGF family proteins.

amino acids 104-155, 171-198

312/330

FIGURE 313

GGGGAGAGGAATTGACCATGTAAAAGGAGACTTTTTTTTTTGGTGGTGGTGGCTGTTGGGTGCCTTGCAAAAAT
 GAAGGATGCAGGACGCAGCTTCTCCTGGAACCGAACGCAATGGATAAACTGATTGTGCAAGAGAGAAGGAAGA
 ACGAAGCTTTTTCTTGTGAGCCCTGGATCTTAACACAAATGTGTATATGTGCACACAGGGAGCATTCAAGAATG
 AAATAAACAGAGTTAGACCCGCGGGGGTGGTGTGTTCTGACATAAATAAATAATCTTAAAGCAGCTGTTCCC
 CTCCCCACCCCCAAAAAAGGATGATTGGAAATGAAGAACCGAGGATTCACAAAGAAAAAAGTATGTTCAATTT
 TTCTCTATAAAGGAGAAAGTGAGCCAAGGAGATATTTTTGGAATGAAAAGTTTGGGGCTTTTTTAGTAAAGTAA
 AGAACTGGTGTGGTGGTGTTCCTTTCTTTTTGAATTTCCACAGAGGAGAGGAAATTAATAATACATCTGC
 AAAGAAATTTAGAGAAGAAAAGTTGACCGCGGCAGATTGAGGCATTGATTGGGGGAGAGAAACCAGCAGAGCA
 CAGTTGGATTTGTGCCTATGTTGACTAAAATTGACGGATAATTGCAGTTGGATTTTCTTCATCAACCTCCTTT
 TTTTTAAATTTTTATTCTTTTGGTATCAAGATCATGCGTTTTCTCTGTTCTTAACCACCTGGATTTCCATCT
 GGATGTTGCTGTGATCAGTCTGAAATACAACCTGTTTGAATTCAGAAGGACCAACACCAGATAAATTATGAATG
 TTGAACAAGATGACCTTACATCCACAGCAGATAATGATAGGTCTAGGTTTAAACAGGGCCCTATTTGACCCCT
 GCTTGTGGTGTGCTGGCTCTTCAACTTCTTGTGGTGGCTGGTCTGGTGCAGGCTCAGACCTGCCCTTCTGTGT
 GCTCCTGCAGCAACCAGTTGAGCAAGGTGATTTGTGTTGCGAAAAACCTGCGTGAGGTTCCGGATGGCATCTCC
 ACCAACACACGGCTGCTGAACCTCCATGAGAACCATAATCCAGATCATCAAAGTGAACAGCTTCAAGCACTTGAG
 GCACTTGGAATCCTACAGTTGAGTAGGAACCATATCAGAACCATTGAAATTGGGGCTTTCAATGGTCTGGCGA
 ACCTCAACACTCTGGAACCTTTGACAATCGTCTTACTACCATCCCGAATGGAGCTTTTGTATACTTGTCTAAA
 CTGAAGGAGCTCTGGTTGCGAAACAACCCCATTGAAAGCATCCCTTCTTATGCTTTTACAGAATTCTTCTTT
 GCGCCGACTAGACTTAGGGGAATTGAAAAGACTTTCATACATCTCAGAAGGTGCCTTTGAAGGTCTGTCCAAC
 TGAGGTATTTGAACCTTGCCATGTGCAACCTTCGGGAAATCCCTAACCTCACACCGCTCATAAACTAGATGAG
 CTGGATCTTTCTGGAATCATTTATCTGCCATCAGGCCTGGCTCTTCCAGGGTTTGATGCACCTTCAAAAAC
 GTGGATGATACAGTCCCAGATTCAAGTGATTGAACGGAATGCCTTTGACAACCTTCAGTCACTAGTGGAGATCA
 ACCTGGCACACAATAATCTAACATTACTGCCTCATGACCTCTTCACTCCCTTGCATCATCTAGAGCGGATACAT
 TTACATCACAACCCTTGGAACGTGAACGTGTGACATACTGTGGCTCAGCTGGTGGATAAAAGACATGGCCCCCTC
 GAACACAGCTTGTTGTGCCCGGTGTAACACTCCTCCCAATCTAAAGGGGAGGTACATTGGAGAGCTCGACCAGA
 ATTACTTCACATGCTATGCTCCGGTGATTGTGGAGCCCCCTGCAGACCTCAATGTCACTGAAGGCATGGCAGCT
 GAGCTGAAATGTCGGGCCTCCACATCCCTGACATCTGTATCTTGGATTACTCCAAATGGAACAGTCATGACACA
 TGGGGCGTACAAAGTGCGGATAGCTGTGCTCAGTGATGGTACGTTAAATTTACAAATGTAACGTGTGCAAGATA
 CAGGCATGTACACATGTATGGTGAGTAATTCGGTTGGGAATACTACTGCTTCAGCCACCCTGAATGTTACTGCA
 GCAACCACTACTCCTTTCTCTTACTTTTCAACCGTCACAGTAGAGACTATGGAACCGTCTCAGGATGAGGCAG
 GACCACAGATAACAATGTGGTCCCCTCCAGTGGTGCAGTGGGAGACCACCAATGTGACCACCTCTCTCACAC
 CACAGAGCACAAGTCGACAGAGAAAACCTTACCATCCAGTGACTGATATAAACAGTGGGATCCCAGGAATT
 GATGAGGTGATGAAGACTACCAAAATCATCATTGGGTGTTTTGTGGCCATCACACTCATGGCTGCAGTGATGCT
 GGTCATTTTCTACAAGATGAGGAAGCAGCACCATCGGCAAAACCATCACGCCCCAACAGGACTGTTGAAATTA
 TTAATGTGGATGATGAGATTACGGGAGACACACCCATGGAAAGCCACCTGCCCATGCCTGCTATCGAGCATGAG
 CACCTAAATCACTATAACTCATACAAATCTCCCTTCAACCACACAACAACAGTTAACACAATAAATTCATACAC
 CAGTTGAGTGCATGAACCGTTATTGATCCGAATGAACCTCTAAAGACAATGTACAAGAGACTCAAATCTAAACAA
 TTTACAGAGTTACAAAAACAAACAATCAAAAAAAGACAGTTTATTAAAAATGACACAAATGACTGGGCTAA
 ATCTACTGTTTCAAAAAAGTGTCTTTACAAAAAACAAGAAAGAAATTTATTTATTAATAAATCTATTG
 TGATCTAAAGCAGACAAAAA

FIGURE 314

MLNKMTLHPQQIMIGPRFNRALFDPLLVLALLQLLVVAGLVRAQTCPSVCSCSNQFSKVIC
VRKNLREVPDGI STNTRLLNLHENQIQIIKVN SFKHLRHLEILQLSRNHIRTIEIGAFNGLA
NLNTLELFDNRLTTIPNGAFVYLSK LKELWLRNNPIESIPSYAFNRIPSLRRLDLGELKRLS
YISEGAFEGLSNLRYLNLAMCNLREIPNLTPLIKLDEL DLSGNHLSAIRPGSFQGLMHLQKL
WMIQSQIQVIERNAFDNLQSLVEINLAHNNLTLLPHDLFTPLHHLERIH LHHNPWNCNCDIL
WLSWWIKDMAPSNTACCARCNTPPNLKGRYIGELDQNYFTCYAPVIVEPPADLNVTEGMAAE
LKCRAS TSLTSVSWITPNGTVMTHGAYKVRIAVLSDGTLNFTNVTVQDTGMYTCMVNSVGN
TTASATLNVTAATTT PFSYFSTVTVETMEPSQDEARTTDNNVGPTPVVDWETTNVTTSLTPQ
STRSTEKTFTIPVTDINSGIPGIDEVMKTTKIIIGCFVAITLMAAVMLVIFYKMRKQHHRQN
HHAPTRTVEIINVDDEITGDTPMESHLPMPAIEHEHLNHYSYKSPFNHTTTVNTINSIHSS
VHEPLLIRMNSKDNVQETQI

Signal sequence:

amino acids 1-44

Transmembrane domain:

amino acids 523-543

N-glycosylation site.amino acids 278-282, 364-368, 390-394, 412-416, 415-419,
434-438, 442-446, 488-492, 606-610**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 183-187

Casein kinase II phosphorylation site.

amino acids 268-272, 417-421, 465-469, 579-583, 620-624

N-myristoylation site.amino acids 40-46, 73-79, 118-124, 191-197, 228-234, 237-243,
391-397, 422-428, 433-439, 531-537

314/330 11/11/94

FIGURE 315

GCGCCGGGAGCCCATCTGCCCCAGGGGCACGGGGCGCGGGGCCGGCTCCCCGCCGGCACAT
 GGCTGCAGCCACCTCGCGCGCACCCCCGAGGCGCCGCGCCCAGCTCGCCCGAGGTCCGTCGGA
 GCGCCCCGGCCGCCCCGGAGCCAAGCAGCAACTGAGCGGGGAAGCGCCCGCGTCCGGGGATC
 GGG**ATG**TCCCTCCTCCTTCTCCTCTTGCTAGTTTCCTACTATGTTGGAACCTTGGGGACTCA
 CACTGAGATCAAGAGAGTGGCAGAGGAAAAGGTCACCTTTGCCCTGCCACCATCAACTGGGGC
 TTCCAGAAAAAGACACTCTGGATATTGAATGGCTGCTCACCGATAATGAAGGGAACCAAAAA
 GTGGTGATCACTTACTCCAGTCGTCTATGTCTACAATAACTTGACTGAGGAACAGAAGGGCCG
 AGTGGCCTTTGCTTCCAATTTCTGAGGAGATGCCTCCTTGACAGATTGAACCTCTGAAGC
 CCAGTGATGAGGGCCGGTACACCTGTAAGGTTAAGAATTCAGGGCGCTACGTGTGGAGCCAT
 GTCATCTTAAAAGTCTTAGTGAGACCATCCAAGCCCAAGTGTGAGTTGGAAGGAGAGCTGAC
 AGAAGGAAGTGACCTGACTTTGCAGTGTGAGTCATCCTCTGGCACAGAGCCCATTTGTGTATT
 ACTGGCAGCGAATCCGAGAGAAAGAGGGAGAGGATGAACGTCTGCCTCCCAAATCTAGGATT
 GACTACAACCACCTGGACGAGTTCTGCTGCAGAATCTTACCATGTCTACTCTGGACTGTA
 CCAGTGCACAGCAGGCAACGAAGCTGGGAAGGAAAGCTGTGTGGTGCGAGTAACTGTACAGT
 ATGTACAAAGCATCGGCATGGTTGCAGGAGCAGTGACAGGCATAGTGGCTGGAGCCCTGCTG
 ATTTTCCTCTTGGTGTGGCTGCTAATCCGAAGGAAAAGACAAAGAAAGATATGAGGAAGAAGA
 GAGACCTAATGAAATTCGAGAAGATGCTGAAGCTCCAAAAGCCCGTCTTGTGAAACCCAGCT
 CCTCTTCCTCAGGCTCTCGGAGCTCACGCTCTGGTTCTTCTCCACTCGCTCCACAGCAAAT
 AGTGCCTCACGCAGCCAGCGGACACTGTCAACTGACGCAGCACCCCAGCCAGGGCTGGCCAC
 CCAGGCATACAGCCTAGTGGGGCCAGAGGTGAGAGGTTCTGAACCAAAGAAAGTCCACCATG
 CTAATCTGACCAAAGCAGAAACCACACCCAGCATGATCCCCAGCCAGAGCAGAGCCTTCCAA
 ACGGTCT**TGA**ATTACAATGGACTTGACTCCCACGCTTTCCCTAGGAGTCAGGGTCTTTGGACTC
 TTCTCGTCATTGGAGCTCAAGTCACCAGCCACACAACCAGATGAGAGGTCATCTAAGTAGCA
 GTGAGCATTGCACGGAACAGATTGAGATGAGCATTTCCTTATACAATACCAAACAAGCAAA
 AGGATGTAAGCTGATTCATCTGTAAAAAGGCATCTTATTGTGCCTTTAGACCAGAGTAAGGG
 AAAGCAGGAGTCCAAATCTATTTGTTGACCAGGACCTGTGGTGAGAAGGTTGGGGAAAGGTG
 AGGTGAATATACCTAAAACCTTTAATGTGGGATATTTTGTATCAGTGCTTTGATTACAAATT
 TTCAAGAGGAAATGGGATGCTGTTTGTAAATTTTCTATGCATTTCTGCAAACCTATTGGATT
 ATTAGTTATTGAGACAGTCAAGCAGAACCCACAGCCTTATTACACCTGTCTACACCATGTAC
 TGAGCTAACCACCTTCTAAGAAACTCCAAAAAAGGAAACATGTGTCTTCTATTCTGACTTAAC
 TTCATTTGTCATAAGGTTTGGATATTAATTTCAAGGGGAGTTGAAATAGTGGGAGATGGAGA
 AGAGTGAATGAGTTTCTCCACTCTATACTAATCTCACTATTTGTATTGAGCCCAAAATAAC
 TATGAAAGGAGACAAAAATTTGTGACAAAGGATTGTGAAGAGCTTTCCATCTTCATGATGTT
 ATGAGGATTGTTGACAAACATTAGAAATATATAATGGAGCAATTGTGGATTTCCTCTCAAAT
 CAGATGCCTCTAAGGACTTTCCTGCTAGATATTTCTGGAAGGAGAAAATACAACATGTCATT
 TATCAACGTCCTTAGAAAGAATTCTTCTAGAGAAAAAGGGATCTAGGAATGCTGAAAGATTA
 CCAACATACCATATAGTCTCTTCTTTCTGAGAAAATGTGAAACCAGAATTGCAAGACTGG
 GTGGACTAGAAAGGGAGATTAGATCAGTTTTCTCTTAATATGTCAAGGAAGGTAGCCGGGCA
 TGGTGCCAGGCACCTGTAGGAAAATCCAGCAGGTGGAGGTTGCAGTGAGCCGAGATTATGCC
 ATTGCACTCCAGCCTGGGTGACAGAGCGGGACTCCGTCTC

316/330

FIGURE 316

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45419

><subunit 1 of 1, 373 aa, 1 stop

><MW: 41281, pI: 8.33, NX(S/T): 3

MSLLLLLLLLVSYYVGTLGTHTEIKRVAEEKVTLPCHHQLGLPEKDTLDIEWLLTDNEGNQKV
VITYSSRHVYNNLTEEQKGRVAFASNFLAGDASLQIEPLKPSDEGRYTCKVKNSGRYVWSHV
ILKVLVRPSKPKCELEGELTEGSDLTLCESSSGTEPIVYYWQRIREKEGEDERLPPKSRID
YNHPGRVLLQNLMTSYSGLYQCTAGNEAGKESCVVRVTVQYVQSIGMVAGAVTGIVAGALLI
FLLVWLLIRRKDKERYEEEEERPNEIREDAEAPKARLVKPSSSSSGSRSSRSGSSSTRSTANS
ASRSQRTLSTDAAPQPGLATQAYSLVGPEVRGSEPKKVHHANLTKAETTPSMIPSQSRAFQTV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 232-251

Downloaded from www.sciencedirect.com

FIGURE 317

CGCGAGGCGCGGGGAGCCTGGGACCAGGAGCGAGAGCCGCTACCTGCAGCCGCCGCCACGGCACGGCAGCCA
 CCATGCGCGCTCCTGTGCTTCGTGCTCCTGTGCGGAGTAGTGGATTTGCGCCAGAAGTTTGTAGTATCACTACT
 CCTGAAGAGATGATTGAAAAAGCCAAAGGGGAACTGCCTATCTGCCATGCAAATTTACGCTTAGTCCCGAAGA
 CCAGGGACCGCTGGACATCGAGTGGCTGATATCACCAGCTGATAATCAGAAGGTGGATCAAGTGATTATTTTAT
 ATTCTGGAGACAAAATTTATGATGACTACTATCCAGATCTGAAAGGCCGAGTACATTTTACGAGTAATGATCTC
 AAATCTGGTGATGCATCAATAAATGTAACGAATTTACAACGTGTGAGATATTGGCACATATCAGTGCAAAGTGAA
 AAAAGCTCCTGGTGTGCAAATAAGAAGATTCACTGGTAGTCTTGTGTTAAGCCTTCAGGTGCGAGATGTTACG
 TTGATGGATCTGAAGAAATTTGGAAGTGACTTTAAGATAAAATGTGAACCAAAGAAGGTTCACTTCCATTACAG
 TATGAGTGGCAAAAATTTGTCTGACTCACAGAAAATGCCACTTCATGGTTAGCAGAAATGACTTCATCTGTTAT
 ATCTGTAAAAAATGCCTCTTCTGAGTACTCTGGGACATACAGCTGTACAGTCAGAAACAGAGTGGGCTCTGATC
 AGTGCCTGTTGCGTCTAAACGTTGTCCCTCCTTCAAATAAAGCTGGACTAATTGCAGGAGCCATTATAGGAACT
 TTGCTTGCTCTAGCGCTCATTGGTCTTATCATCTTTTGTGTCGTAAGGCGCAGAGAAGAAAAATATGAAAA
 GGAAGTTCATCAGATATCAGGGAAGATGTGCCACCTCCAAAGAGCCGTACGTCCACTGCCAGAAGCTACATCG
 GCAGTAATCATTATCCCTGGGGTCCATGTCTCCTTCCAACATGGAAGGATATTCCAAGACTCAGTATAACCAA
 GTACCAAGTGAAGACTTTGAACGCACTCCTCAGAGTCCGAACTCTCCACCTGCTAAGTTCAAGTACCCTTACAA
 GACTGATGGAATTACAGTTGTATTAATATGGACTACTGAAGAATCTGAAGTATTGTATTATTTGACTTTATTTT
 AGGCCTCTAGTAAAGACTTAAATGTTTTTAAAAAAGCACAAAGGCACAGAGATTAGAGCAGCTGTAAGAACAC
 ATCTACTTTTATGCAATGGCATTAGACATGTAAGTCAGATGTGATGTCAAATTTAGTACGAGCCAAATCTTTGT
 TAAAAAACCTATGTATAGTGACACTGATAGTTAAAAGATGTTTTATTATATTTTCAATAACTACCACTAACAA
 ATTTTAACTTTTTCATATGCATATTTCTGATATGTGGTCTTTTAGGAAAAGTATGGTTAATAGTTGATTTTCAA
 AGGAAATTTTAAATTTCTTACGTTCTGTTTAAATGTTTTGCTATTTAGTTAAATACATTGAAGGGAAATACCCG
 TTCTTTTCCCCTTTTATGCACACAACAGAAACACGCGTTGTGATGCCTCAAACATTTTTTTATTTGCAACTACA
 TGATTTTACACAATTTCTTTAAACAACGACATAAAATAGATTTTCTTGTATATAAATAACTTACATACGCTCCA
 TAAAGTAAATTTCTCAAAGGTGCTAGAACAAATCGTCCACTTCTACAGTGTTCTCGTATCCAACAGAGTTGATGC
 ACAATATATAAATACTCAAGTCCAATATTAATAAATTTAGGCACTTGACTAACTTTAATAAATTTCTCAAACCTA
 TATCAATATCTAAAGTGCAATATATTTTTTAAAGAAAGATTATTTCTCAATAACTTCTATAAAAAATAAGTTTGATGG
 TTTGGCCCATCTAACTTCACTACTATTAGTAAGAAGTTTAACTTTTAAATGTGTAGTAAGGTTTATTTCTACCTT
 TTTCTCAACATGACACCAACACAATCAAAAACGAAGTTAGTGAGGTGCTAACATGTGAGGATTAATCCAGTGAT
 TCCGGTCACAATGCATTCCAGGAGGAGGTACCCATGTCACTGGAATTGGGCGATATGGTTTTATTTTTCTTCCC
 TGATTTGGATAACCAAATGGAACAGGAGGAGGATAGTGATTTCTGATGGCCATTCCCTCGATACATTCCTGGCTT
 TTTTCTGGGCAAAGGTTGCCACATTGGAAGAGGTGGAATATAAGTTCTGAAATCTGTAGGGAAGAGAACACAT
 TAAGTTAATTTCAAAGGAAAAAATCATCATCTATGTTCCAGATTTCTCATTAAAGACAAAGTTACCCACAACACT
 GAGATCACATCTAAGTGACACTCCTATTGTGAGGTCTAAATACATTAAAAACCTCATGTGTAATAGGCGTATAA
 TGTATAACAGGTGACCAATGTTTTCTGAATGCATAAAGAAATGAATAAACTCAAACACAGTACTTCTTAAACAA
 CTTCAACCAAAAAAGACCAAAACATGGAACGAATGGAAGCTTGTAAGGACATGCTTGTTTTAGTCCAGTGGTTT
 CCACAGCTGGCTAAGCCAGGAGTCACTTGGAGGCTTTTAAATACAAAACATTGGAGCTGGAGGCCATTATCCTT
 AGCAAATAATGCAGAAACAGAAAATCAACTACCGCATGTTCTCACTTATAAGTGGGAGGTAATGATAAGAACT
 TATGAACACAAAGAAGGAAACAATAGACATTGGAGTCTATTTGAGAGGGGAGGTTGGGAGAAGGAAAAGGAGCA
 GAAAAGATAACTATTGAGTACTGCCCTTACACCTGGGTGATGAAATAATATGTACAACAAATCCCTGTGACACA
 TGTTTACCTATGGAACAAACCTTCATGTGTATCCCTAAACCTAAAATAAAAGTTAAAAAARAAAAA
 AA
 AAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 318

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA82361

><subunit 1 of 1, 352 aa, 1 stop

><MW: 38938, pI: 7.86, NX(S/T): 3

MALLLCFVLLCGVVDFAFARSLSITTPPEEMIEKAKGETAYLPCKFTLSPEDQGPLDIEWLISPA
DNQKVDQVIILYSGDKIYDDYYPDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQCKVKK
APGVANKKIHVLVVLVKPSGARCXYVDGSEEIGSDFKIKCEPKESLPLQYEWQKLSDSQKMPT
SWLAEMTSSVISVKNASSEYSGTYSCVRNRVGSQCLLRNLNVPPSNKAGLIAGAIIGTLL
ALALIGLIIIFCCRKKRREEKYEKEVHHDIREDVPPPKSRTSTARSYIGSNHSSLGSMSPSNM
EGYSKTQYNQVPSEDFERTPQSPTLPPAKFKYPYKTDGITVV

Signal sequence.

amino acids 1-19

Transmembrane domain:

amino acids 236-257

N-glycosylation sites.

amino acids 106-110, 201-205, 298-302

Tyrosine kinase phosphorylation sites.

amino acids 31-39, 78-85, 262-270

N-myristoylation sites.

amino acids 116-122, 208-214, 219-225, 237-243, 241-247,
245-251, 296-302

Myelin P0 protein.

amino acids 96-125

FIGURE 320

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA88002

><subunit 1 of 1, 206 aa, 1 stop

><MW: 23799, pI: 9.12, NX(S/T): 3

MNFQQRLQSLWTLARPFPCPLLATASQMOMVVLPCLGFTLLLWSQVSGAQGQEFHFGPCQVK
GVVPQKLWEAFWAVKDTMQAQDNITSARLLQQEVLQNVSDAESCYLEVHTLLEFYLKTVFKNH
HNRTVEVRTLKSFSTLANNFVLIVSQLQPSQENEMFSIRDSAHRRFLLFRRAFKQLDVEAAL
TKALGEVDILLTWMQKFYKL

Signal sequence:

amino acids 1-42

N-glycosylation sites.

amino acids 85-89, 99-103, 126-130

FIGURE 321

AAGGAGCAGCCCGCAAGCACCAAGTGAGAGGCATGAAGTTACAGTGTGTTTCCCTTTGGCTC
CTGGGTACAATACTGATATTGTGCTCAGTAGACAACCACGGTCTCAGGAGATGTCTGATTTC
CACAGACATGCACCATATAGAAGAGAGTTTCCAAGAAATCAAAGAGCCATCCAAGCTAAGG
ACACCTTCCCAAATGTCACTATCCTGTCCACATTGGGAGACTCTGCAGATCATTAAGCCCTTA
GATGTGTGCTGCGTGACCAAGAACCTCCTGGCGTTCTACGTGGACAGGGTGTTCAAGGATCA
TCAGGAGCCAAACCCCAAATCTTGAGAAAAATCAGCAGCATTGCCAACTCTTTCCTCTACA
TGCAGAAAACCTCTGCGGCAATGTCAGGAACAGAGGCAGTGTCACTGCAGGCAGGAAGCCACC
AATGCCACCAGAGTCATCCATGACAACCTATGATCAGCTGGAGGTCCACGCTGCTGCCATTAA
ATCCCTGGGAGAGCTCGACGTCTTTCTAGCCTGGATTAATAAGAATCATGAAGTAATGTTCT
CAGCTTGATGACAAGGAACCTGTATAGTGATCCAGGGATGAACACCCCCTGTGCGGTTTACT
GTGGGAGACAGCCCACCTTGAAGGGGAAGGAGATGGGGAAGGCCCTTGCAGCTGAAAGTCC
CACTGGCTGGCCTCAGGCTGTCTTATTCCGCTTGAAAATAGGCAAAAAGTCTACTGTGGTAT
TTGTAATAAACTCTATCTGCTGAAAGGGCCTGCAGGCCATCCTGGGAGTAAAGGGCTGCCTT
CCCATCTAATTTATTGTAAAGTCATATAGTCCATGTCTGTGATGTGAGCCAAGTGATATCCT
GTAGTACACATTGTACTGAGTGGTTTTTCTGAATAAATTCATATTTTACCTATGA

FIGURE 322

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA92282
><subunit 1 of 1, 177 aa, 1 stop
><MW: 20452, pI: 8.00, NX(S/T): 2
MKLQCVSLWLLGTILILCSVDNHGLRRLISTDMHHIEESFQEIKRAIQAKDTFPNVTILST
LETLQIIKPLDVCCVTKNLLAFYVDRVFKDHQEPNPKILRKISSIANSEFLYMQKTLRQCQEQ
RQCHCRQEATNATRVIHNDNYDQLEVHAAAIKSLGELDVFLAWINKNHEVMFSA

Signal sequence:

amino acids 1-18

N-glycosylation sites.

amino acids 56-60, 135-139

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 102-106

N-myristoylation site.

amino acids 24-30

Actinin-type actin-binding domain signature 1.

amino acids 159-169

FIGURE 323

CCCGTGCCAAGAGTGACGTAAGTACCGCCTATAGAGTCTATAGGCCCACTTGGCTTCGTTAG
AACGCGGCTACAATTAATACATAACCTTATGTATCATAACATACGATTTAGGTGACACTAT
AGAATAACATCCACTTTGCCTTTCTCTCCACAGGTGTCCACTCCCAGGTCCAACCTGCACCTC
GGTTCTATCGATAATCTCAGCACCAGCCACTCAGAGCAGGGCACG**ATG**TTGGGGGCCCCGCT
CAGGCTCTGGGTCTGTGCCTTGTGCAGCGTCTGCAGCATGAGCGTCCTCAGAGCCTATCCCA
ATGCCTCCCCACTGCTCGGCTCCAGCTGGGGTGGCTGATCCACCTGTACACAGCCACAGCC
AGGAACAGCTACCACCTGCAGATCCACAAGAATGGCCATGTGGATGGCGCACCCCATCAGAC
CATCTACAGTGCCCTGATGATCAGATCAGAGGATGCTGGCTTTGTGGTGATTACAGGTGTGA
TGAGCAGAAGATACCTCTGCATGGATTTTCAGAGGCAACATTTTGGATCACACTATTTTCGAC
CCGGAGAACTGCAGGTTCCAACACCAGACGCTGGAAAACGGGTACGACGTCTACCACTCTCC
TCAGTATCACTTCCTGGTCAGTCTGGGCCGGGCGAAGAGAGCCTTCCTGCCAGGCATGAACC
CACCCCCGTACTCCCAGTTCCTGTCCCGGAGGAACGAGATCCCCCTAATTCACCTCAACACC
CCCATACACGGCGGCACACCCGGAGCGCCGAGGACGACTCGGAGCGGGACCCCTGAACGT
GCTGAAGCCCCGGGCCCCGGATGACCCCGGCCCCGGCCTCCTGTTACAGGAGCTCCCGAGCG
CCGAGGACAACAGCCCGATGGCCAGTGACCCATTAGGGGTGGTCAGGGGCGGTTCGAGTGAAC
ACGCACGCTGGGGGAACGGGGCCCGGAAGGCTGCCGCCCTTCGCCAAGTTCATC**TAG**GGTCTG
CTGG

FIGURE 324

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA142238

><subunit 1 of 1, 251 aa, 1 stop

><MW: 27954, pI: 9.22, NX(S/T): 1

MLGARLRLWVCALCSVCSMSVLRAYPNASPLLGSWSGGLIHLYTATARNSYHLQIHKNGHVD
GAPHQTIYSALMIRSEDAGFVVITGVMSRRYLCMDFRGNI FGSHYFDPENCRFQHQTLENGY
DVIHSPQYHFLVSLGRAKRAFLPGMNPPPYSQLSRNEIPLIHFNTPIPRRHTRSAEDDSE
RDPLNLVKPRARMTPAPASCSQELPSAEDNSPMASDPLGVVRGGRVNT HAGGTGPEGCRPFA
KFI

Important features of the protein:**Signal peptide:**

amino acids 1-24

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 175-179

N-myristoylation site.

amino acids 33-39, 100-106, 225-231, 229-235

HBGF/FGF family proteins

amino acids 73-124

FIGURE 326

MGCLWGLALPLFFFCWEVGVSGSSAGPSTRRADTAMTTDDTEVPAMTLAPGHAALETQTL
 ETSSRASTPAGPIPEAETRGAKRISPARETRSF TKTS PNF MVLIATSVETSAASGSPEGAGM
 TTVQTITGSDPEEAI FDTLCTDD SSEEAKTLTMDILT LAHTSTEAKGLSSESSASSDGPHPV
 ITPSRASESSASSDGPHPVITPSRASESSASSDGPHPVITPSWSPGSDVTLLAEALVTVTNI
 EVINCSITEIETTTSSIPGASDIDLIPTEGVKASSTSDPPALPDSTEAKPHITEVTASAETL
 STAGTTESAAPHATVGTPLPTNSATEREVTAPGATTL SGALVTVSRNPLEETSALSVETPSY
 VKVSGAAPVSIEAGSAVGKTTSFAGSSASSYSPSEAALKNF TPSETPTMDIATKGPFPTSRD
 PLPSVPPTTTNSSRGTNSTLAKITTS AKTTMKPQQPRRLPGRGRPQT

N-glycosylation sites:

amino acids 252-256, 445-449, 451-455

cAMP-and cGMP-dependent protein kinase phosphorylation site.

amino acids 84-90

Casein kinase II phosphorylation sites.

amino acids 37-41, 108-112, 131-135, 133-137, 148-152, 165-169,
 246-250, 254-258, 256-260, 269-273, 283-287, 333-337, 335-339,
 404-408, 414-418, 431-435

N-myristoylation sites.

amino acids 2-8, 19-25, 117-123, 121-127, 232-238, 278-284, 314-
 320, 349-355, 386-392, 397-403, 449-455

ATP/GTP-binding site motif A (P-loop).

amino acids 385-393

FIGURE 328

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45410
><subunit 1 of 1, 431 aa, 1 stop
><MW: 46810, pI: 6.45, NX(S/T): 6
MFFGGEGSLTYTLVIICFLTLRLSASQNCLKKSLEDVVIDIQSSLKSGIRGNEPVYTSTQED
CINSCCSTKNISGDKACNLMIFDTRKTARQPNCYLFFCPNEEACPLKPAKGLMSYRIITDFP
SLTRNLPSQELPQEDSLLHGQFSQAVTPLAHHHTDYSKPTDISWRDTLSQKFGSSDHLEKLF
KMDEASAQLLAYKEKGHSQSSQFSSDQEIAHLLPENVSALPATVAVASPHHTSATPKPATLL
PTNASVTPSGTSQPQLATTAPPVTTVTSPPTTLISTVFTRAAATLQAMATTAVLTTTFQAP
TDSKGSLETIPFTEISNLTNLNTGNVYNPTALSMSNVESSTMNKTASWEGREASPGSSSQGSV
PENQYGLPFEKWLLIGSLLFGVLFLVIGLVLLGRILSESLRRKRYSRLDYLINGIYVDI
```

Signal sequence.

amino acids 1-25

Transmembrane domain.

amino acids 384-405

N-glycosylation sites.

amino acids 72-76, 222-226, 251-255, 327-331, 352-356

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 415-419

Tyrosine kinase phosphorylation site.

amino acids 50-57

N-myristoylation sites.

amino acids 4-10, 48-54, 315-321

Downloaded from www.sagepub.com at 11:11 11 November 2009

FIGURE 329

CTCCACGGTGTCCAGCGCCAGAA**ATG**CGGCTTCTGGTCCTGCTATGGGGTTGCCTGCTGCT
 CCCAGGTTATGAAGCCCTGGAGGGGCCAGAGGAAATCAGCGGGTTCGAAGGGGACACTGTGT
 CCCTGCAGTGCACCTACAGGGAAGAGCTGAGGGACCACCGGAAGTACTGGTGCAGGAAGGGT
 GGGATCCTCTTCTCTCGCTGCTCTGGCACCATCTATGCAGAAGAAGAAGGCCAGGAGACAAT
 GAAGGGCAGGGTGTCCATCCGTGACAGCCGCCAGGAGCTCTCGCTCATTGTGACCCTGTGGA
 ACCTCACCCCTGCAAGACGCTGGGGAGTACTGGTGTGGGGTCGAAAAACGGGGCCCCGATGAG
 TCTTTACTGATCTCTCTGTTTCGTCTTTCCAGGACCCTGCTGTCTCCCTCCCTTCTCCCAC
 CTTCCAGCCTCTGGCTACAACACGCCTGCAGCCCAAGGCAAAAAGCTCAGCAAACCCAGCCCC
 CAGGATTGACTTCTCTGGGCTCTACCCGGCAGCCACCACAGCCAAGCAGGGGAAGACAGGG
 GCTGAGGCCCCCTCATTGCCAGGGACTTCCCAGTACGGGCACGAAAGGACTTCTCAGTACAC
 AGGAACCTCTCCTCACCCAGCGACCTCTCCTCCTGCAGGGAGCTCCCGCCCCCATGCAGC
 TGGACTCCACCTCAGCAGAGGACACCAGTCCAGCTCTCAGCAGTGGCAGCTCTAAGCCCAGG
 GTGTCCATCCCGATGGTCCGCATACTGGCCCCAGTCTGTGTGCTGCTGAGCCTTCTGTGAGC
 CGCAGGCCTGATCGCCTTCTGCAGCCACCTGCTCCTGTGGAGAAAGGAAGCTCAACAGGCCA
 CGGAGACACAGAGGAACGAGAAGTTCTGGCTCTCACGCTTGACTGCGGAGGAAAAGGAAGCC
 CTTTCCAGGCCCTGAGGGGGACGTGATCTCGATGCCTCCCTCCACACATCTGAGGAGGA
 GCTGGGCTTCTCGAAGTTTGTCTCAGCG**TAG**GGCAGGAGGGCCCTCCTGGCCAGGCCAGCAGT
 GAAGCAGTATGGCTGGCTGGATCAGCACCGATTCCCGAAAGCTTTCACCTCAGCCTCAGAG
 TCCAGCTGCCCCGACTCCAGGGCTCTCCCCACCCTCCCCAGGCTCTCCTCTTGATGTTCCA
 GCCTGACCTAGAAGCGTTTGTGAGCCCTGGAGCCCAGAGCGGTGGCCTTGCTCTTCCGGCTG
 GAGACTGGGACATCCCTGATAGGTTACATCCCTGGGCAGAGTACCAGGCTGCTGACCCTCA
 GCAGGGCCAGACAAGGCTCAGTGGATCTGGTCTGAGTTTCAATCTGCCAGGAACCTTGGGC
 CTCATGCCCAGTGTGCGACCCTGCCTTCTCCCACTCCAGACCCACCTTGTCTTCCCTCCC
 TGGCGTCTCAGACTTAGTCCCACGGTCTCCTGCATCAGCTGGTGTGATGAAGAGGAGCATGCT
 GGGGTGAGACTGGGATTCTGGCTTCTCTTTGAACCACCTGCATCCAGCCCTTCCAGGAAGCCT
 GTGAAAAACGTGATTCTGGCCCCACCAAGACCCACCAAAACCATCTCTGGGCTTGGTGCAG
 GACTCTGAATTCTAACAATGCCCAGTGAAGTGTGCACTTGAGTTTGAGGGCCAGTGGGCCTG
 ATGAACGCTCACACCCCTTCAGCTTAGAGTCTGCATTTGGGCTGTGACGTCTCCACCTGCCC
 CAATAGATCTGCTCTGTCTGCGACACCAGATCCACGTGGGGACTCCCTGAGGCCTGCTAAG
 TCCAGGCCTTGGTCAGGTGAGTGCACATTGCAGGATAAGCCCAGGACCGGCACAGAAGTGG
 TTGCCTTTNCCATTTGCCCTCCCTGGNCCATGCCTTCTTGCTTTGGAAAAAATGATGAAGA
 AAACCTTGGCTCCTTCTTGTCTGGAAAGGGTTACTTGCCTATGGGTTCTGGTGGCTAGAGA
 GAAAAGTAGAAAACCAGAGTGCACGTAGGTGTCTAACACAGAGGAGAGTAGGAACAGGGCGG
 ATACCTGAAGGTGACTCCGAGTCCAGCCCCCTGGAGAAGGGGTCGGGGGTGGTGGTAAAGTA
 GCACAACACTACTATTTTTTTTCTTTTTTCCATTATTATTGTTTTTTAAGACAGAATCTCGTGCT
 GCTGCCCAGGCTGGAGTGCAGTGGCACGATCTGCAAACCTCCGCCTCCTGGGTTCAAGTGATT
 CTTCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGCACGCACCACCACACCTGGCTAATT
 TTTGTACTTTTAGTAGAGATGGGGTTTACCATTGTTGGCCAGGCTGGTCTTGAACCTCTGAC
 CTCAAATGAGCCTCCTGCTTCAGTCTCCCAAATTGCCGGGATTACAGGCATGAGCCACTGTG
 TCTGGCCCTATTTCTTTAAAAAGTGAAATTAAGAGTTGTTTCAGTATGCAAAACTTGGAAAG
 ATGGAGGAGAAAAAGAAAAGGAAGAAAAAAATGTACCCATAGTCTCACCAGAGACTATCAT
 TATTTTCGTTTTGTGTACTTCTTCCACTCTTTTCTTCTTACATAATTTGCCGGTGTTCTT
 TTTACAGAGCAATTATCTTGTATATAACAATTTGTATCCTGCCTTTTCCACCTTATCGTTCC
 ATCACTTTATTCCAGCACTTCTCTGTGTTTTACAGACCTTTTTTATAAATAAAATGTTTCATCA
 GCTGCATAAAAAAAAAAAAAA

FIGURE 330

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44196

<subunit 1 of 1, 332 aa, 1 stop

<MW: 36143, pI: 5.89, NX(S/T): 1

MRLLVLLWGCLLLPGYEALGPEEISGFEGDTVSLQCTYREELRDHRKYWCRKGGILFSRCS
GTIYAEEEGQETMKGRVSIRDSRQELSLIVTLWNLTLQDAGEYWCGVEKRGPDSELLISLFV
FPGPCPPSPSPPTFQPLATTRLQPKAKAQQTQPPGLTSPGLYPAATTAKQGKTGAEAPPLPG
TSQYGHERTSQYTGTSPPHATSPAGSSRPPMQLDSTSAEDTSPALSSGSSKPRVSIPMVRI
LAPVLVLLSLLSAAGLIAFCSHLLLRKEAQQATETQRNEKFWLSRLTAEKEAPSQAPEGD
VISMPPLHTSEEELGFSKFVSA

Important features:**Signal peptide:**

amino acids 1-17

Transmembrane domain:

amino acids 248-269

N-glycosylation site.

amino acids 96-99

Fibrinogen beta and gamma chains C-terminal domain.

amino acids 104-113

Ig like V-type domain:

amino acids 13-128